Fields Institute Communications
Volume 1, 1993

Gauge Theory of the Falling Cat

Richard Montgomery

June 8, 1993

1. Introduction and Summary

Kane and Scher [18] proposed a mechanical model in order to explain and better understand how a falling cat rights herself. Their model cat consists of two identical axi-symmetric rigid bodies which are joined by a special 'no-twist' joint. The first problem is for the model cat to right herself while in freefall with no angular momentum beginning from an upside-down position. Kane and Scher, and earlier Rademaker and ter Braak [25], proposed a specific strategy for doing this. But they did not study the problem of finding the general strategy for performing the flip. A second problem is for the to perform her trick in an optimal way. These can be viewed as problems in control theory.

In earlier papers [22], [21] we developed a general theory for the attitude, or orientation control, of deformable bodies in freefall with zero angular momentum. These papers were outgrowths of work by Wilczek and Shapere [27] and Guichardet [12] . The main point of these earlier works is that a dictionary can be developed between the gauge theory of the physicist's and mathematicians, and the problems occuring in the orientation control of deformable bodies. Briefly, in this dictionary the space of shapes of the body plays the role of the base space, or space-time in the physicist's gauge theory. Its tangent space is the space of controls. The state space, or configuration space of the body, is principal bundle of the theory. The gauge group is the group of rigid reorientations of the body. The gauge field summarizes the condition that the angular momentum be zero.

The purpose of the present paper is to apply our general theory to the Kane-Scher cat. Without the special no-twist joint, the shape space is the group $SO(3)$ with an element in it representing the attitude of one half of the cat relative to a frame fixed to the other. The configuration space is $Q = SO(3) \times SO(3)$ with one $SO(3)$ for each body half, and the gauge group is $SO(3)$, acting diagonally on the configuration space.

Here are our main results.

- The shape space of the model cat is the real projective plane $\mathbb{R}P^2$ embedded in $SO(3) \cong \mathbb{R}P^3$ as a (projective) linear subspace.
- The collision states in which the two body halves coincide form the line at infinity, $\mathbb{R}P^1 \subset \mathbb{R}P^2$.
- The reachable states starting from any given state of Q form an $SO(3)$ embedded anti-diagonally in Q. The projection from the reachable states to the

This paper is in final form and no version will be submitted for publication elsewhere.

no-twist shape space forms the ($\mathbf{Z}_2$-quotient of the) Hopf bundle which is a principal $O(2)$ bundle.

- The control system is defined by an axial vector potential or gauge field –one which changes sign under passing to "the other side" of shape space.
- By deleting the line at infinity shape space becomes the affine plane and the structure group reduces from $O(2)$ to $SO(2)$.
- **In particular, the differential equations which must be solved to calculate the reorientation induced by any control strategy can be reduced to a single quadrature.**
- The optimal control problem is equivalent to the equations of motion of an (axially-) charged particle travelling on the projective plane under the influence of an axially symmetric (axial) magnetic field and axially symmetric metric.
- We present a simplifying feedback transformation, linear in the controls and induced by a coordinate transformation of shape space, which maps the control system to the maximally symmetric system of this type on $SO(3)$,i.e. the one in which all moments of inertias are equal and the joint is at the two body's center of mass.
- If the metric, or cost, on shape space is the pull-back of the rotationally invariant metric under the change of variables which induce the feedback transformation, then the optimal loops are the original loops of Kane and Scher.

Kane and Scher calculated the reorientation, or holonomy, suffered when their model cat traversed a particular class of loops in its shape space. These are the loops described in the final item above, and consist of one body half, say the back, describing a circular cone relative to a frame attached to the other. Such a loop is a geometric circle in the projective plane with its usual rotationally invariant metric. In performing their calculation they made a particular nice choice of gauge, or local section, for the full bundle $Q = SO(3) \times SO(3) \to SO(3)$. We will see how their gauge is suggested naturally from the group theory of the situation. It is essential to our calculations as well.

2. Acknowledgements

I would like to thank Michael Enos for many valuable conversations. In particular, the idea of using the maximally symmetric cost function on the asymmetric Kane-Scher cat is in large part due to him. I would also like to thank Shankar Sastry, Richard Murray, Zexiang Li and Jerrold Marsden for useful comments. This work was partly sponsored by NSF grant DMS-8807219.

3. The General Setting

3.1. A Gauge Theory for Deformable Bodies. We begin by recalling the relations which have been discovered between the theory of orientation control of a deformable body and gauge theory. This will also help to set notation. Some of the original references are [27], [28], [30] [12], [21], [22], [12].

Imagine that we are in freefall with zero total angular momentum. Our problem is to reorient ourselves, say right-side-up, by changing our shape. Such a problem is faced by gymnasts, falling cats, satellites, etcetera. Our first objective is to control our net orientation. This can be represented by a rotation matrix g. Our control variables are the deformations dx of our shape. Our possible shapes are represented (locally) by a continuous vector variable x.

Remark 1. *Shape deformations are in turn implemented by torques or linear forces applied to joints. In this paper we ignore the problem of how to choose these generalized forces in order to obtain the velocity controls dx needed for a desired reorientation. This problem is sometimes called the 'dynamical problem'. Thus the full problem divides into two parts, the kinematic which we solve, and the dynamic which we don't. For more on the dynamic problem see Bloch et al [4]. From examples one finds that the 'dynamic problem' can often be solved by essentially differentiating the velocities and inverting a matrix to solve for the forces necessary. In these cases the kinematic problem is the central problem. (G. Walsh, private communication.)*

Let Q denote the configuration space of a given deformable body. We will assume that the motion of the body's center of mass motion is fixed as it is in free-fall. This effectively gets rid of the translation subgroup of the group of rigid motions of space. The group G of rotations about the center of mass remains, and acts on Q by rigidly rotating configurations. We write this action as

$$q \in Q \longmapsto gq = q' \in Q, \tag{1}$$

where $g \in G$ is the rotation. Two configurations have the same shape if and only if they are related in this way by some rotation. Thus the *shape space* S is the quotient space:

$$S = Q/G$$

whose points consist of G-orbits. Let

$$\pi : Q \to S$$

denote the map which assigns to each configuration q its shape $x = \pi(q)$. We say that G acts freely if $gq = q$ (for some q) implies that $g = e$, the identity of the group. In this case S is a smooth manifold and π gives Q the structure of a principal G-bundle. We recall the definition:

Definition 1. *$\pi : Q \to S$ is a principal G-bundle if there is a covering of S by open sets $U \subset S$ together with a family of diffeomorphisms*

$$\phi_U : U \times G \to \pi^{-1}(U) \subset Q,$$

$$\phi_U : (x, g) \longmapsto q$$

called "local trivializations" with the property that if $q = \phi_U(x, g)$ and $g_1 \in G$ then $g_1 q = \phi_U(x, g_1 g)$.

Definition 2. *A local section of $\pi : Q \to S$ is a smooth map $s : U \subset S \to Q$ satisfying $\pi(s(x)) = x$*

Every local trivialization uniquely determines a local section

$$s : U \subset S \to Q$$

by the formula

$$s(x) = \phi_U(x, e)$$

Conversely, every local section uniquely determines a local trivialization by the formula

$$\phi_U(x, g) = g s(x)$$

We may think of local sections as smooth local choices of origin (identity) for the fibers $Q_x = \pi^{-1}(x) \sim G$. A local section $s(x)$ is a realization of the shape x as an actual configuration $q = s(x)$ in inertial space.

Recall that by listing the columns of an orthogonal matrix g we obtain an orthonormal frame. In this manner, we may think of the map $q = gs(x) \to g$ as a choice of frame for each configuration q with $\pi(q) \in U$. In the physics literature local sections are often referred to as local gauges.

The control system we will be using is equivalent to conservation of angular momentum. We can think of angular momentum as a vector-valued differential one-form $M(q)dq$ on the configuration space Q. For each deformation δq of $q \in Q$ the angular momentum yields a vector $M(q)\delta q$ which represents the corresponding total angular momentum resulting from this deformation of q. In terms of a local trivialiation we have the formula:

$$M(q)dq = g(I(x)g^{-1}dg + m(x)dx) \tag{2}$$

Here $I(x)$ is the locked inertia tensor for the configuration $s(x)$. This means it is the moment of inertia tensor of the effective rigid body obtained by locking all of the body's joints in the shape x when it is oriented in space according to $s(x)$. The configuration change $s(x) \to s(x + dx)$ results in a total angular momentum $m(x)dx$. The form $g^{-1}dg$ is the angular velocity with respect to a "body" fixed frame. ("Body" is in quotes because this frame depends on the choice of gauge $s(x)$.) More precisely, $g^{-1}dg$ is the pull-back of the right-invariant Maurer-Cartan form on G induced by the local trivialiation: $\pi^{-1}(U) \subset Q \to U \times G \to G$. All of these one-forms take values in $Lie(SO(3))$, the Lie algebra of 3×3 skew-symmetric matrices. $Lie(SO(3))$ represents the space of angular velocities and is naturally isomorphic to $\mathbb{R}^3$. Under this isomorphism the Lie bracket (commutator) of matrices turns into the cross-product of vectors and the action of $SO(3)$ on $Lie(SO(3))$ by conjugation ($\omega \to g\omega g^{-1}$) becomes the usual action of $SO(3)$ on $\mathbb{R}^3$ ($\omega \to g\omega$).

Remark 2. *Set $I(x,g) = gI(x)g^{-1}$ and $m(x,g) = gm(x)dxg^{-1}$. These are the locked inertia tensor and total angular momentum form of the configuration $\phi_U(x,g) = gs(x)$. And $gI(x)^{-1}m(x)dxg^{-1} = I(x,g)^{-1}m(x,g)$.*

We assume that the body's total angular momentum is constant. This will be true if the body is in free-fall and air friction is negligible. The control system can then be written down once the value of the initial angular momentum is known. We will assume that the angular momentum is zero. Left-multiplying our expression (2) for total angular momentum by $gI^{-1}g^{-1}$ we obtain

$$dg + gI(x)^{-1}m(x)dx = 0. \tag{3}$$

This is our control system. Rewritten in the standard control notation it is

$$\begin{aligned} \dot{x} &= u \\ \dot{g} &= -g\Gamma(x)(u) \end{aligned} \tag{4}$$

where u, the control, is a tangent vector to the shape space S and where

$$\Gamma(x) = I(x)^{-1}m(x) : T_xS \to Lie(SO(3)). \tag{5}$$

Shapere and Wilczek [28] refer to this formula as "the master gauge" and we will occassionally refer to it as such. The formula can be found in Guichardet's paper [12]. It is implicit in Smale's work on topology and mechanics in the sense that he has a construction which when carried out concretely lead to this formula.

Remark 3. *If the angular momentum is some nonzero value μ, then the control system is*

$$dg = -g\Gamma(x)dx + gI(x)^{-1}g^{-1}\mu$$

The last term is a drift term. It breaks the symmetry group from $SO(3)$ (for the case $\mu = 0$) to the subgroup of rotations about the μ axis.

3.1.1. Terminology. The vector-valued one-form Γ above will be refered to as the connection one-form. Shapere and Wilczek [27] refer to it as the master gauge field. Marsden refers to it as the 'natural mechanical connection'. The formula for Γ together with its relation to gauge theory was first made explicit by Guichardet [12].

In the context of the geometry of principal bundles our control system (4) is called the equation of parallel translation. It is a time-dependent linear differential equation for g whose coefficients depend on the curve $x(t)$ in shape space.

The one-form $(dg + g\Gamma(x))g^{-1}$ is the expression relative to a local trivialization Φ_U of a one-form defined on all of Q. This one-form is also called the connection one-form. Q is a left principal bundle. **Warning** Most differential geometry texts,eg Kobayashi-Nomizu [19] use right-principal bundles, in which case this formula becomes $g^{-1}dg + g^{-1}\Gamma g$.)

3.2. Optimal Control and SubRiemannian Geometry. Kinetic energy defines a Riemannian metric on Q. As an exercise, the reader can check that the zero angular momentum deformations are exactly those tangent vectors to Q which are orthogonal in to all infinitesimal rotations. (See [22].) As a consequence we may and will identify the space of zero angular momentum deformations at q with our control space at q, namely the tangent space to shape space S at $\pi(q)$. The space of controls at q thus inherits an inner product, or positive definite quadratic form, namely the restriction of the kinetic energy to

$$\mathcal{D}_q = ker(M(q)dq) = \text{ zero-angular momentum deformations of q} \tag{6}$$

More generally, let $\langle u, u\rangle_x$ be a positive definite quadratic form on the space of controls $u \in T_xS$ at x which depends smoothly on $x \in S$; in other words, a Riemannian metric on the shape space. We will investigate the optimal control problem which it defines for our control law.

Problem: Minimize

$$E = \int_0^T \frac{1}{2}\langle u(t), u(t)\rangle_{x(t)} dt$$

among all controls u which steer the state q from an intitial state $q_0 \in Q$ to a final state $q_1 \in Q$ in time T under our control system. (Recall that in a local trivialization $q = (x, g)$.)

How do we come up with this metric on S? We just described one possibility, the one induced by kinetic energy on the configuration space. The choice made by Shapere and Wilczek was dictated by the desire to minimize power expenditure due to transfer of energy to the surrounding fluid. We will take the point of view that the fundamental object is the control system and the fundamental problem is providing an algorithm for simply getting to the desired point (steering) in a computationally feasible manner. **From this point of view then, a good metric is one for which the optimal control problem is computationally simple.** Such a metric is found at the end of this paper.

Whatever the choice of metric, this problem is a special case of the general problem of finding minimizing subRiemannian geodesics. Recall that a distribution on a manifold Q is a smooth subbundle $\mathcal{D} \subset TQ$.

Definition 3. *A subRiemannian structure on a manifold Q consists of a distribution $\mathcal{D}$ on Q together with a smoothly varying positive inner product $\langle\cdot,\cdot\rangle_q$ defined on each plane $\mathcal{D}_q$ of this distribution.*

Remark 4. *Our deformable body problems have additional structure beyond that of the subRiemannian metric. Namely the subRiemannian structures we discussed admit $G = SO(3)$ as an isometry group and this isometry group acts transverse to the distribution. Consequently $\langle\cdot,\cdot\rangle$ and $\mathcal{D}$ are projectable by π to S. In other words $d\pi_q(\mathcal{D}_q) = T_xS$ and $\langle v, w\rangle_q = \langle d\pi_q v, d\pi_q w\rangle_x$ whenever $\pi(q) = x$ and $v, w \in \mathcal{D}_q$.*

Associated to any subRiemannian metric we have two additional optimal control problems. All three problems are minimization problems on the space of absolutely continuous paths joining q_0 to q_1.

- Minimize the length $\int_0^T \sqrt{\langle u(t), u(t)\rangle_{q(t)}}dt$
- Minimize the time T of travel between q_0 and q_1, subject to the constraint that $u(t) \in \mathcal{D}_{q(t)}$ and the bounds $\langle u(t), u(t)\rangle \leq 1$.

It is well-known to experts that these three minimization problems are equivalent and that their corresponding value functions are related just as in Riemannian geometry.

3.3. Hamilton's Equations and Magnetic Fields. We will review the basic facts concerning the Hamiltonian equations governing the solutions to any of the three above optimal control problems. For details we refer the reader to [22] [11] [26] [23].

Let $X_1, \cdots X_m$ be an orthonormal frame field for $\mathcal{D}$ relative to the given inner product. Let $P_i : T^*Q \to \mathbf{R}$ be the corresponding momentum functions:

$$P_i(q,p) = p(X_i(q)).$$

($p \in T_q^*Q$ and $X_i(q) \in T_qQ$ so they naturally pair together to form a number.) Then the Hamiltonian which governs the **normal** optimal controls is

$$H_n = \frac{1}{2}\Sigma_{i=1}^m P_i^2.$$

It is easy to see that this function is independent of choice of frame. In particular, it is globally defined, even though the P_i may only be locally defined. (For the definitions of normal vs. abnormal minimizers see L.C. Young [32] or the recent paper [23].)

With respect to this frame our control system is written

$$\dot{q} = \Sigma u_i(t)X_i(q(t)).$$

Hamilton's equations imply

$$u_i(t) = P_i$$

and

$$\dot{u}_i = \Sigma\Omega_{ij}P_j$$

where Ω is the skew-symmetric matrix with entries $\Omega_{ij} = \{P_i, P_j\}$, the Poisson brackets of the momentum functions. See [26] for a derivation of these observations. The entries Ω_{ij} are in turn the momentum functions for the vector field $-[X_i, X_j]$ obtained from Lie bracket, and can be thought of as a kind of 'curvature' of the distribution. These equations generalize the equations of motion of a particle in a magnetic field.

The Hamiltonian system for the **abnormal** extremals is the Dirac type system with constraints

$$P_i = 0$$

and abnormal Hamiltonian

$$H = \Sigma u_i P_i$$

where the u_i are the optimal controls. The corresponding integral curves are precisely the characteristics for the annihilator of $\mathcal{D}$, which is a submanifold of T^*Q. To our knowledge, this was first observed by Hsu [17]. A proof can also be found in [23].

Theorem 1. *Any optimal control $u(t)$ for the minimum energy, length, or time problem induces a curve $q(t)$ along which there exists a continuous costate $p(t)$ such that $(q(t), p(t))$ solves Hamilton's differential equations corresponding to either H_n or H_{ab}. If the extremal is normal (for H_n) then this solution curve is smooth.*

Open Problem Show that the minimizing curves $q(t)$ are smooth in the abnormal case.

The previous theorem characterizes minimizers. But do they exist? We recall the classical conditions and theorem of Chow.

Definition 4. *The distribution $\mathcal{D}$ is bracket generating at q if it admits a frame $E_i, i = 1, 2, \ldots, r$ such that the E_i together with their iterated Lie brackets, $[E_i, E_j]$, $[E_i, [E_j, E_k]], \ldots$ span T_qQ upon evaluation at q . It is bracket generating if it is bracket generating at all points q.*

The bracket generating property is independent of choice of frame. It is also generic. Chow's theorem says that if $\mathcal{D}$ is bracket generating and Q is connected then any two points of Q can be joined by a curve tangent to $\mathcal{D}$, that is, by a solution to the control system. (If the distribution is analytic then the bracket generating condition is also a necessary condition for joining any two points.) Combining this with the Arzela-Ascoli theorem we have:

Theorem 2. *Suppose that $\mathcal{D}$ is bracket generating and Q is connected. Also suppose that Q is compact or that the X_i are complete or that the sR metric is the restriction to $\mathcal{D}$ of a complete Riemannian metric on Q. Then any two points of Q can be joined by a solution to the optimal control problem with these as endpoints, i.e. by a subRiemannian geodesic which is minimizing.*

4. Effectively Planar Deformable Bodies as Charged Particles

We return to our class of examples. In the next section we will show that the Kane-Scher model cat satisfies the following properties. Property (A): Shape space is two-dimensional. Property (B): The structure group G is one-dimensional. (It is $O(2)$.) We can then find coordinates (x, y, z) on the configuration space for which the first two coordinates coordinatize shape space and the last coordinate is an angle representing the group direction. The gauge field has the form $dz + \Gamma$ where $\Gamma = A_1(x, y)dx + A_2(x, y)dy$. Finally, we can alway arrange that the coordinates (x, y) are such that the metric on shape space is diagonal

$$\langle \cdot, \cdot \rangle = d^2s = E(x, y)dx^2 + G(x, y)dy^2.$$

The normal Hamiltonian is

$$H_n = \frac{1}{2}\{\frac{1}{E}(p_1 - A_1(x, y)p_3)^2 + \frac{1}{G}(p_2 - A_2(x, y)p_3)^2\}.$$

Since the variable z does not occur explicitly we have

$$\dot{p}_3 = 0$$

If we interpret the constant of motion p_3 as an electric charge then the normal Pontrjagin Hamiltonian H_n is the Hamiltonian for a particle with this charge and a unit mass travelling on the Riemannian surface S through the magnetic field whose vector potential is Γ. The magnetic field is the (pseudo) scalar field $B = \frac{d\Gamma}{\text{area form}}$.

Any abnormal extremal must lie in the zero -level set of the magnetic field [[22]], [[24]]. For the falling cat we will show that this set is empty so that we need not worry about the abnormal extremals.

There is a subtlety which occurs in the falling cat example concerning the difference between $SO(2)$ and $O(2)$ gauge fields. The essence of this subtlety is that for $O(2)$ gauge fields, the 'magnetic field', which is now the curvature of an $O(2)$ gauge field need not be a two-form in the standard sense. Instead, it is a two-form with values in some real line bundle (the adjoint bundle) over S. For the falling cat S turns out to be the real projective plane and this line bundle is the 'orientation bundle'.

To appreciate how this subtlety comes about it is best to consider our optimal control problem when Q is a general principal G-bundle, G an arbitrary Lie group. The optimal control Hamiltonian H_n is a G invariant function on T^*Q. It follows that the Hamiltonian flow descends to $(T^*Q)/G$. Using the connection we can identify this quotient with $T^*S \oplus Ad^*(Q)$, the direct sum of the cotangent bundle of shape space with the co-adjoint bundle, which is the vector bundle with fiber $Lie(G)^*$ associated to Q by the coadjoint action. Roughly speaking the elements in the co-adjoint bundle are the Lagrange multipliers which enforce the nonholonomic constraint. Now we can identify T^*S with TS using the metric. The curvature F of the connection-form is a two-form with values in the adjoint bundle $Ad(Q)$ dual to Ad^*Q. With respect to these reduced variables and identifications the Euler-Lagrange equations which govern the normal optimal extremals are

$$\nabla_{\dot{x}}\dot{x} = \mu F(\dot{x}, \cdot)$$

and

$$\frac{D\mu}{dt} = 0$$

These are equations for a curve (x, μ) with $\mu(t) \in Ad^*_{x(t)}Q$ in in the coadjoint bundle. ∇ is the Levi-Civita connection on S so that if the fiber variable μ were 0 the first equation would say that x is a geodesic on S. The right-hand side of the first equation defines a one-form along x which we identify it with a vector field along x by using the metric.These are "Wong's" equations [31] for the motion of a particle in the Yang-Mills field Γ over S. For more on this see [22].

5. Flipping the Model Cats of Kane and Scher

5.1. The Model and the Gauge of Kane and Scher. The Kane-Scher model [18] cat consists of two identical axially symmetric rigid bodies, called the front and back halves, joined together along their symmetry axes by a special type of joint. These symmetry axes represent the cat's backbone.

See Figure 1.

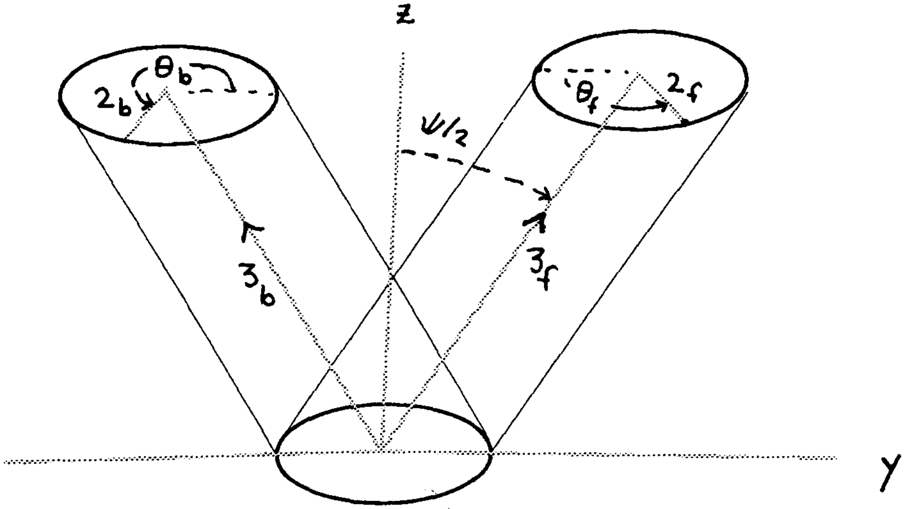


Figure 1

For purposes of visualization we think of each body as a right circular cylinder. We label the cylinders f and b for "front" and "back".

We will begin our analysis by supposing the joint to be ball-and-socket; that is, there will be no constraint on the relative motion of the two halves other than that they are joined at this joint. Later we will impose the no-twist constraint of Kane-Scher. This constraint is holonomic. Imposing it is equivalent to replacing the ball-and-socket joint with a special type of joint meant to account for the more limited class of relative deformations allowed between two vertebrae.

The angle between the two symmetry axes, 3_f and 3_b, will be denoted by ψ. 3_f and 3_f are oriented so that they each point **out of** the common joint. Mark a point on the surface of each body half (cylinder) and connect this point to the symmetry axis by a vector orthogonal to the symmetry axis. These vectors represent the cat's legs. Label them 2_f (for front) and 2_b (for back). They are principal directions of inertia for their body half. Let $\mathcal{P}$ denote the plane in 3 -space containing the symmetry axes 3_f and 3_b. Let θ_f be the angle between 2_f and this plane $\mathcal{P}$ and θ_b the corresponding angle for 2_b. More information is required to uniquely specify these angles. We insist that they increase as their corresponding leg rotates in a positive sense about its symmetry axes. And we suppose that when the angle ψ is between 0 and π and the angles θ_f and θ_b are 0 that the components of the feet vectors in the direction of the perpindicular bisector $3_f + 3_b$ of the symmetry axes is positive. After choosing a local section, this bisector will represent the direction 'up'. So, we are saying that this configuration corresponds to feet pointing up and body bent upwards. With these conventions, together with continuity, the cat's shape is completely specified by the coordinates $(\psi, \theta_f, \theta_b)$. **However there are coordinate singularities at** $\psi = 0$ and $\psi = \pi$. For in these cases the symmetry axes are collinear and so do not determine a plane. Also the coordinates $(\pi, \theta_f, \theta_b)$ and $(\pi, \theta_f + \theta_0, \theta_b + \theta_0)$ represent the same shape for any angle θ_0.

We are now going to describe a specific configuration $\sigma(\psi, \theta_f, \theta_b)$ which **realizes** the shape with coordinates $(\psi, \theta_f, \theta_b)$. **In other words, σ will be a local section for the bundle $Q \to S$ = shape space.** To dot this fix an inertial system of axes, xyz. (See §1.) We require that the plane $\mathcal{P}$ is the yz plane, and that the bisector $(3_f + 3_b)$ of the angle formed by the symmetry axes is pointed along the y-axis. The y-axis represents the up direction. $\psi/2$ is then the angle between each symmetry axes and the y-axis. These requirements, together with continuity, uniquely specify σ. For instance $\sigma(\pi, 0, 0)$ is the configuration in which the cat's backbone lies on the z axis and its legs are pointing straight up. We take this to be the cat's initial configuration. It represents the initially held upside down cat. The 4-tuple $(\psi, \theta_f, \theta_b, g) \in S^1 \times S^1 \times S^1 \times SO(3)$ corresponds to the configuration $gs(\psi, \theta_f, \theta_b)$. These 4-tuples define the Kane-Scher coordinates on the configuration space $Q = SO(3) \times SO(3)$ of the Kane-Scher cat .

By slight abuse of notation we will write $\sigma(\psi)$ for the curve $\sigma(\psi, 0, 0)$. Let $R(\theta_f, \theta_b)$ denote the two-parameter group of **material** symmetries of the model cat obtained by rotating each body half about its symmetry axes by the given angle:

$$R(\theta_f, \theta_b)(g_f, g_b) = (g_f exp(\theta_f \mathbf{e_3}), g_b exp(\theta_b \mathbf{e_3}).$$

Notation. If ω is a vector in space then $exp(\omega)$ denotes the operation of rotation about the ω axes by $\|\omega\|$ radians. If we identify vectors with skew-symmetric matrices in the standard way (ω corresponds to the skew symmetric operator $\mathbf{x} \longmapsto \omega \times \mathbf{x}$) then this is the usual exponential of a matrix. Thus $exp(\theta \mathbf{e_x})$ is a rotation about the inertial z-axis by θ radians.

The Kane-Scher coordinates $(\psi, \theta_f, \theta_b, g)$ can alternatively be defined by

$$q(g, \psi, \theta_f, \theta_b) = gR(\theta_f, \theta_b)\sigma(\psi) \tag{7}$$

This last equation illustrates the group theoretic signifiance of their frame. The configuration space for the model cat is

$$Q = SO(3)_f \times SO(3)_b$$

The bodies are such that the isometry group of Q is

$$Isom(Q) = SO(3) \times SO(2)_f \times SO(2)_b \times_s \mathbb{Z}_2$$

The first factor represents spatial rotations. The last three terms are the material symmetries corresponding to rotating the front body about its symmetry axis, rotating the back body about its axis, and switching the two body halves $(\mathbb{Z}_2)$. The identity component of Isom(Q) is of course

$$Isom(Q)^0 = SO(3) \times SO(2)_f \times SO(2)_b$$

and (7) describes its action. $\sigma(\psi)$ is a slice to the action of Isom $(Q)^0$. This just means that it intersects each orbit once. The subscript "s" in front of the $\mathbb{Z}_2$ factor stands for semidirect product. It accounts for the fact that when we switch the two bodies the actions of $SO(2)_f$ and $SO(2)_b$ must also be switched.

5.2. The Connection Form : Ball-and-Socket Case. Our control system is $dg + g\Gamma dx = 0$ where Γ is given by the "master formula"

$$\Gamma = I^{-1}m \tag{8}$$

described in §2. We now calculate I, m and Γ with respect to the Kane-Scher frame. Figure 8 summarizes this frame (choice of gauge; local section) s.

From Figure 1 we see that

$$m = I_3[(se_3 + ce_2)d\theta_f + (se_3 - ce_2)(-d\theta_b)]$$

is the angular momentum due to the deformation $\sigma(\psi, \theta_f, \theta_b) \longmapsto \sigma(\psi_1 + d\psi, \theta_f + d\theta_f, \theta_b d\theta_b)$. Here I_3 is the moment of inertia of either body held about its symmetry axes and

$$s = sin(\frac{\psi}{2}), \quad c = cos(\frac{\psi}{2})$$

A more involved but still straight forward calculation shows:

$$I = 2diag(I_1 + ml^2c^2, I_1s^2 + I_3c^2 + ml^2s^2, I_1c^2 + I_3s^2)$$

is the inertia tensor, locked at $\sigma(\psi, \theta_f, \theta_b)$. Here $I_1 = I_2$ are the equal moments of inertia of a body half when the corresponding axis passes through the body's center of mass, m is the total mass of a body half and l is the distance of its center of mass from the joint.

Remark 5. *It is clear from Figure 1 that the Kane-Scher frame diagonalizes I. Up to the labelling of axes this property characterizes the Kane-Scher frame.*

Define dimensionless parameters, α, β, by using I_1 as a unit of measure

$$I_3 = \alpha I_1 \ ; \ ml^2 = \beta I_1$$

A direct calculation using the above results yields

$$\Gamma = (\Phi_+(\psi)\hat{e}_2 + \Phi_-(\psi)\hat{e}_3)d\theta_f + (-\Phi_+(\psi)\hat{e}_2 + \Phi_-(\psi)\hat{e}_3)(-d\theta_b) \tag{9}$$

where

$$\Phi_+ = \frac{1}{2}\,\frac{\alpha c}{s^2 + \alpha c^2 + \beta s^2}$$

$$\Phi_- = \frac{1}{2}\,\frac{\alpha s}{c^2 + \alpha s^2}$$

Note that the $\Phi_\pm$ satisfy the symmetry properties $\Phi_+(-\psi) = \Phi_+(\psi)$, $\Phi_-(-\psi) = -\Phi_-(\psi)$.

Remark 6. *The form (9) for Γ above and the symmetry properties of $\Phi_\pm$ follows directly from the symmetries of the kinetic energy and the group theoretic properties of the Kane-Scher frame as described in the previous section. In the language of gauge theory the Φ's are the Higgs fields corresponding to the material symmetries.*

5.3. No Twist. Each cat half has its own angular velocity vector $\omega_{\mathbf{f}}, \omega_{\mathbf{b}}$, as viewed from an inertial frame. Consider the components of these vectors along the respective half's symmetry axis, $\omega_{3,f} = \omega_f \cdot e_{3f}$, $\omega_{3,b} = \omega_b \cdot e_{3b}$. Kane and Scher introduced the no-twist constraint:

$$\omega_{3,f} = -\omega_{3,b}$$

as a way of modelling the cat's joint. The quantities $\omega_{3,f}$ and $\omega_{3,b}$ are called the spins. They are the Noether-conserved quantities or momentum maps corresponding to the rotational symmetry about these axes. In Kane-Scher coordinates we have $\omega_{3,f} = d\theta_f$ and $\omega_{3,b} = d\theta_b$, the no-twist constraint reads $d\theta_f = -d\theta_b$, and so is a holonomic constraint:

$$\theta_f = -\theta_b + constant.$$

We take the constant to be zero. Thus no-twist shape space is coordinatized by (ψ, θ) where

$$\theta = \theta_f = -\theta_b.$$

Roughly, this constraint says that the cat cannot break her own back. As observed by Mike Enos, we can think of the constraint as saying that the two body halves are identical tin cans joined so that they roll without slipping along their common lids.

Consider our choice of gauge $\sigma(\psi, \theta_f, \theta_b)$ restricted to the no-twist subspace $\theta_f = -\theta_b$. By abuse of notation we write it as

$$\sigma(\psi, \theta) = \sigma(\psi, \theta, -\theta).$$

Figure 1 should convince the reader that any change in ψ alone is a zero-angular momentum deformation: $m(\sigma)(\frac{d\sigma}{d\psi}) = 0$. On the other hand, a change in θ alone leads, by the symmetry of the figure, to a net angular momentum parallel to the 3 (equals z) axis. Now any deformation of no-twist shape space is a linear superposition of these two deformations and consequently can only have angular momentum along the 3 axis. Again by symmetry, the locked inertia tensor $I(\sigma(\psi, \theta))$ is diagonal with respect to the xyz axes. **It follows that the connection one form, $\sigma^*\Gamma = I^{-1}m(\sigma, \cdot)$ has only one non-vanishing component and this is in the 3 direction. It follows that the model cat can only rotate about the z-axis!**

From this last result we see that the configuration, or more precisely, the reachable set of the zero-angular momentum no-twist model cat can be coordinatized by variables (ψ, θ, χ) according to the rule To summarize

$$(\psi, \theta, \chi) \longleftrightarrow g(\chi)R(\theta, \theta)c(\psi) \tag{10}$$

Here

$$g = g(\chi) = exp(\chi \mathbf{e_z})$$

is a rotation about the z-axis by χ radians.

We now find the explicit form for the connection. To do this, plug the no-twist constraint into the general form (9) of the connection form. This yields

$$\Gamma = \Gamma(\psi)\mathbf{e_3}d\theta; \quad \Gamma(\psi) = \frac{\alpha s}{c^2 + \alpha s^2}\mathrm{e}_3. \tag{11}$$

(Please excuse the double use of "Γ". It should cause no confusion.) It is the no-twist connection one-form and describes the effect of no-twist deformations on the model cat's orientation. Explicitly, for **small** loops c we have the reorientation formula

$$g_1 = e^{\chi \mathbf{e}_3} g_0$$

where g_0 and g_1 are the initial and final orientations and where

$$\chi = -\oint_c \Gamma(\psi) d\theta.$$

We will say precisely what we mean by "small" later.

There are two remarkable things concerning formula (11). The first we have already pointed out, but is worth repeating.**Fact 1:** The reorientation can be obtained by a single quadrature. (this is false for the general parallel transport law $\dot{g} = A(t)g$ on SO(3) and hence for the model cat with built with a ball-and-socket joint.) **Fact 2:** The parameter β which describes the distance between the joint and either body's center of mass does not occur in the formula. It follows that the reorientation of the model cat is independent of this distance and in particular we would obtain the same reorientation χ even if the bodies were joined at their mass centers (provided the ratio α of moments of inertia is kept the same).

5.4. The Global Structure of the No-twist Constraint. We have just seen the remarkable fact that the no-twist connection form takes values in a one-dimensional subalgebra of $Lie(SO(3))$ and its as a consequence the fact that its holonomy group lie in a one-parameter subgroup of $SO(3)$. At first glance, one might think this group is $SO(2)$. But in fact it is $O(2)$. In the language of gauge theory, imposing the no-twist constraint has reduced the structure group from $SO(3)$ to $O(2) \subset SO(3)$. The following cartoon illustrates a holonomy in $O(2)$ but not in $SO(2)$. In this cartoon the body halves must pass through each other at the top of their motion's arc.

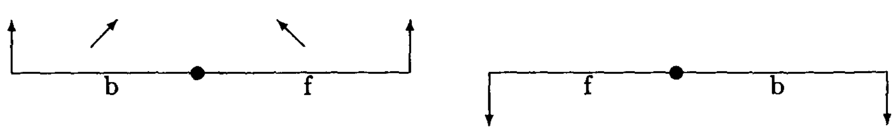


Figure 2

From the previous section we know that the set Q_{nt} of configurations accessible by zero angular momentum no twist deformations from a given configuration is locally coordinatized by (ψ, θ, χ). We have a corresponding no-twist shape space and bundle:

$$\pi : Q_{nt} \to S_{nt}$$

We will now show that this bundle is the principal $O(2)$-bundle

$$SO(3) \to \mathbb{R}P^2$$

In order to do this it is convenient to reorient the symmetry axis of the back body in the opposite sense of our original orientation. This defines new coordinates on Q. (This is the choice made by Kane and Scher.) Also the angle between the symmetry axes with this new orientation is related to the old orientation by $\psi_{new} = \psi_{old} + \pi$. In terms of our old Kane-Scher coordinates this coordinate change is given by

$$(g, \psi, \theta_f, \theta_b) \to (g, \psi + \pi, \theta_f, -\theta_b) = (g, \phi, \theta_1, \theta_2)$$

With respect to these new coordinates the no-twist constraint is

$$d\theta_1 = d\theta_2$$

And our slice $\sigma(\psi)$ becomes

$$\sigma_{new}(\phi) = (e^{\frac{\phi}{2}\mathbf{e}_1}, e^{-\frac{\phi}{2}\mathbf{e}_1})$$

so that the no-twist configurations are coordinatized by (χ, ϕ, θ) with

$$g = e^{\chi_1 e_3}$$

$$\phi = \psi + \pi$$

$$\theta = \theta_1 = \theta_2$$

Write $(g_1, g_2) \in Q$ as

$$g_1 = e^{\chi_1 \mathbf{e}_3} e^{\phi_1 \mathbf{e}_1} e^{\theta_1 \mathbf{e}_3}$$

$$g_2 = e^{\chi_2 \mathbf{e}_3} e^{\phi_2 \mathbf{e}_1} e^{\theta_2 \mathbf{e}_3}$$

(Euler coordinates). Then the no-twist configuration space $Q_{nt} \subset Q$ is defined by

$$Q_{nt} := \begin{cases} \chi_1 &= & \chi_2 & = & \chi \\ \phi_1 &= & -\phi_2 & := & \psi/2 \\ \theta_1 &= & \theta_2 & = & \theta \end{cases}$$

The projection $\pi : Q \to S$ for full configuration space can written

$$\pi(g_1, g_2) = g_1^{-1} g_2.$$

For points in Q_{nt} this reads

$$\pi(g_1, g_2) = e^{-\theta e_3} e^{\phi e_1} e^{\theta e_3} = exp(\phi(cos\theta e_3 + sin\theta e_2))$$

Thus $\pi(Q_{nt}) = exp(\mathbb{R}^2)$ A simple calculation now shows that

$$exp(\mathbb{R}^2) = \mathbb{R}P^2 \subset \mathbb{R}P^3 = SO(3) = exp(\mathbb{R}^3).$$

This shows that the no-twist shape space S_{nt} is the real projective plane as claimed in the introduction.

We will now show that $Q_{nt}$is isomorphic, as a smooth $O(2)$-bundle over S_{nt}, to the frame bundle $SO(3))$ of $\mathbb{R}P^2$. Set

$$\varepsilon_3 = exp(\pi e_3)$$

and consider the involution:

$$i(g_1, g_2) = (\varepsilon_3 g_2 \varepsilon_3, \varepsilon_3 g_1 \varepsilon_3)$$

of Q. A straightforward calculation shows that

$$\varepsilon_3 e^{\chi e_3} e^{\phi e_1} e^{\theta e_3} \varepsilon_3 = e^{\chi e_3} e^{-\phi e_1} e^{\theta e_3}$$

so that in our Euler angle coordinates

$$i(\chi_1, \phi_1, \theta_1, \chi_2, \phi_2, \theta_2) = (\chi_2, -\phi_2, \theta_2, \chi_1, -\phi_1, \theta_1).$$

This shows that the fixed point set of i is Q_{nt}. Observe that

$$i(g_1, \varepsilon_3 g \varepsilon_3) = (g, \varepsilon_3 g_1 \varepsilon_3)$$

and that any element g_2 of G can be written in the form $\varepsilon_3 g \varepsilon_3$. This shows that the fixed point set of i is the set elements $(g_1, g_2) \in Q = G \times G$ of the form $(g, \varepsilon_3 g \varepsilon_3)$. This proves that $Q_{nt} \simeq SO(3)$. Define the map $\pi : Q_{nt} = SO(3) \to S_{nt} = \mathbb{R}P^2$projection $Q \to S = Q/G$ to $Q_{nt} \subset Q$. From the work we have done, we see that π is the quotient map for the action of the $O(2) \subset G$. It follows that π is a version of the standard Hopf fibration $S^3 = SU(2) \to S^2$, namely a quotient of it by $\mathbb{Z}_2$.

Now i is an isometric involution with respect to the kinetic energy metric <u>or</u> the subRiemannian structure of Q. It follows that Q_{nt} is totally geodesic with respect to both the Riemanian and subRiemannian structures. Consequently any free motion or optimal path for Q, whose end points line on Q_{nt} must lie entirely in Q_{nt}. (These facts can also be proved using the dynamical invariance of the spins ω_3, f, ω_3, b under the corresponding Hamiltonian flows.) **In physical terms no torques are needed to impose the no-twist joint beyond those needed to impose the connection of the two halves (i.e. to impose a ball-and-socket joint) as long as initial velocities are tangent to the no-twist configuration space** , i.e., as long as the cat does not start off by trying to break her own back.

5.4.1. No-Twist Symmetries. The symmetry group of the no-twist configuration space consists of those elements of the full isometry group which take no-twist configurations to other no-twist configurations. This group is

$$Isom(Q_{nt}) = SO(2) \times SO(2) \times_s \mathbb{Z}_2$$

where the circles $SO(2)$ are rotations about the 3-axis but are not the previous material symmetry rotations, but rather a "diagonal" combination. Specifically, $(R_l, R_r, 1) \in SO(2) \times SO(2) \times_s \mathbb{Z}_2$ in the identity component of the group acts by

$$(R_l, R_r)(g_1, g_2) = (R_l g_1 R_r^{-1}, R_l g_2 R_r^{-1})$$

(The $\mathbb{Z}_2$ switches the body halves as before.) The last two factors act nontrivially on the no-twist shape space $\mathbb{R}P_2$ with the elements R_r of the second circle factor acting by a standard rotation of the projective plane about a fixed point. This fixed point is the fully stretched out state. We will mark it and consider it to be the origin of the projective plane. That is, it is the center of the standard affine chart.

At the other extreme are the collision states. These are the shapes where the two body halves coincide, that is, the cat is completely folded up. Such shapes form the 'line at infinity' in the projective plane. Recall that this line is in fact a topological circle ($\mathbb{R}P^1$) and that one can parameterize this circle as the angle between the front and back legs of the completely folded up model cat.

From the general theory developed in §3 we know that the normal optimal extremals for the no-twist cat are characterized as the motions of a charged particle on the projective plane travelling through a magnetic field. The metric on projective space which defines the cost function should be taken to have the same symmetries. For example, the metric induced from the kinetic energy on Q has this property. As we discussed, there is one hitch. The particle is "axial" in the sense that it is associated to an $O(2)$ instead of a $U(1) = SO(2)$ gauge theory. But with this proviso, we can say that the optimal curves are the motions of charged particles travelling through a rotationally symmetric magnetic field on the projective plane.

5.4.2. Global properties of the Curvature. Recall the coordinate formula

$$\Gamma = \frac{\alpha s}{c^2 + \alpha s^2} d\theta$$

for the connection form. Its curvature is

$$\Omega = dA = \alpha \frac{1 - (\alpha - 1)s^2}{(c^2 + \alpha s^2)^2} ds \wedge d\theta$$

where, recall that $s = sin\psi/2$ so that $ds = \frac{1}{2}cd\psi$. Observe that

$$1 - (\alpha - 1)s^2 > 0 \text{ so that } \Omega \neq 0.$$

This is because $1 \geq \alpha - 1$ or $2 \geq \alpha$ which in turn follows from the inequality $I_1 + I_2 \geq I_3$, valid for the eigenvalues of any inertia tensor. ($2 = \alpha$ corresponding to a degenerate planar body.)

Since Ω is never zero there can be no abnormal extremals. See §4. Consequently we have shown:

Theorem 3. *All optimal controls (minimizing geodesics) for the Kane-Scher model cat are normal, no matter what cost function (metric on shape space) is chosen.*

On the other hand there is something paradoxical about the fact that Ω is never 0. $\mathbb{R}P^2$ is not orientable so it does not admit any nonvanishing two-forms! How do we resolve this contridiction? The resolution is that the structure group of the bundle is $O(2)$, not $SO(2)$ and consequently the curvature is not an honest two-form. Instead, under a change of frame $\sigma \mapsto \varepsilon\sigma$ corresponding to an element ε of $O(2)$ which represents the nontrivial class in $O(2)/SO(2) = \mathbb{Z}_2$ we have

$$\Omega \mapsto \varepsilon\Omega\varepsilon = -\Omega$$

Such nontrivial representatives ε occur in the bundle transitions from the usual affine chart (the disc's interior) to one intersecting the line at infinity (one containing collision shapes).

We can now say what we meant by "small" for the loop c in the reconstruction (holonomy) formula following [9]. "Small" simply means that the loop is contractible in $\mathbb{R}P^2$. (Recall that $\pi^1(\mathbb{R}P^2) = \mathbb{Z}_2$ so that there is only one other type of loop a besides contractible one.) Any loop for which the two bodies do not collide, i.e. does not intersect the line at infinity is small.

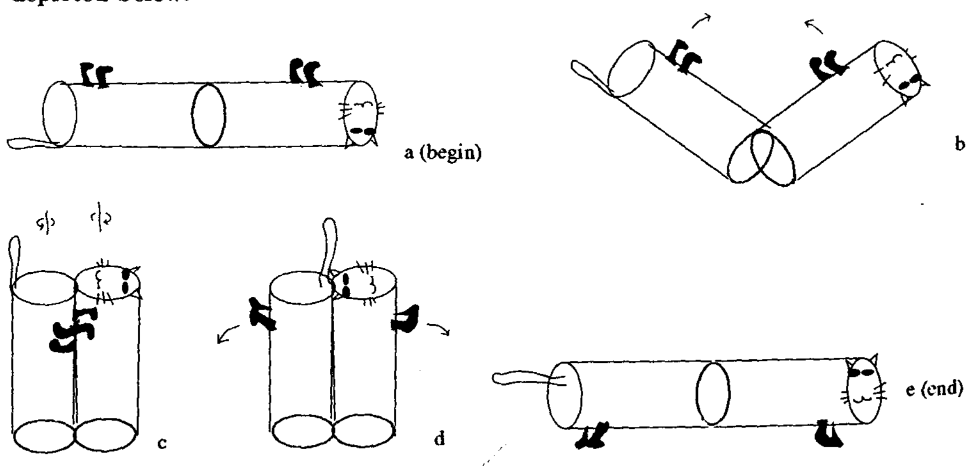
5.5. Summary. We summarize the observations of §4 with a theorem

Theorem 4. *The no-twist condition of Kane and Scher defines totally geodesic submanifolds, $Q_{nt} \subset Q$, which we call the no-twist configuration space , and $S_{nt} \subset S$, which we call the no-twist shape space. These are totally geodesic in both the Riemannian (free motion) and subRiemannian (controlled motion) senses. In physical terms this means that when undergoing the optimally controlled or free motions, no torques are necessary in order to keep the joint of no-twist type. The torques imposed and the motions undergone in such motions are exactly the same as if the joint were replaced by a spherical (ball and socket) joint and the corresponding problem (free or optimally controlled) solved.*

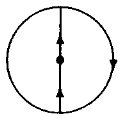
Q_{nt} is diffeomorphic to $SO(3)$ and hence to $\mathbb{R}P^3$. S_{nt} is diffeomorphic to the real projective plane $\mathbb{R}P^2$. The line at infinity in the plane corresponds to shapes where the two body halves are coincident (complete collisions). The restriction of the projection $Q \to S$ to Q_{nt} gives it the structure of a principal $O(2)$-bundle over S_{nt}. This bundle is isomorphic to the Hopf fibration bundle $SO(3) \to \mathbb{R}P^2$. The no-twist connection, Γ, on this bundle, which is defined by the vanishing of angular momentum, has symmetry group $SO(2) \times_s \mathbb{Z}_2$ (in addition to the symmetry under the action of the structure group which every connection enjoys) where the rotation acts on $\mathbb{R}P^2$ in the standard manner by rotating about an axis. The metric on no-twist shape space induced by kinetic energy also has this as an isometry group. Consequently, both the geodesic and the optimal control flows on Q_{nt} are completely integrable. The latter is equivalent to the motion of an $O(2)$- charged particle moving on the projective plane under the influence of the axially symetric magnetic field which is equal to the curvature of the no-twist connection.

6. Some Specific Reorientations and Steering Strategies

6.1. A Cartoon. Probably the simplest path resulting in the cat flip is the one depicted below.



In the disc model of the projective plane the curve corresponding to this cartoon is



6.2. A simple flip. For illustrative purposes we have drawn the collision shapes (b) and (c) so that the bodies do not coincide. If we are honest to the model they should coincide and then configurations (b) and (c) would represent the same shapes, but with one rotated by $e^{pi\mathbf{e}_2}$ with respect to the other. This rotation represents the the non-identity component of the fiber $O(2)$ of $SO(3) \to I\!RP^2$. It is amusing to note that either of the two pieces of this curve on the projective plane represent the non-trivial element of $\pi_1(I\!RP^2)$ and together they show that $\pi_1(I\!RP^2) = \mathbb{Z}_2$.

This reorientation maneuver is universal in the sense that it works no matter what the values of the parameters α or β are. It violates Kane and Scher's small back bending criteria (criterion (c) at the beginning of [18]) rather spectacularly.

6.3. The Rademaker-terBraak Meridians. A special class of motions, apparently first studied by Rademaker and ter Braak [25] (see also [18]) are those along the meridians $\psi=$ const, $0 \leq \theta \leq 2\pi$ (see figure). The resulting reorientation is

$$\begin{aligned}\chi(\psi_0) &= \int_0^{2\pi} \phi(\psi_0) d\theta \\ &= \frac{\alpha s}{c^2 + \alpha s^2} 2\pi\end{aligned}$$

If we set $\chi = \pi$ corresponding to the cat flipping we obtain

$$\frac{1}{2} = \frac{\alpha s}{c^2 + \alpha s^2}$$

or

$$\psi_0 = 2sin^{-1}\left(\frac{\alpha \pm \sqrt{1 - \alpha + \alpha^2}}{\alpha - 1}\right)$$

6.4. The Conical Motions of Kane and Scher. A slightly more general class of loops in shape space are the conical motions introduced by Kane and Scher [18]. These are closed curves in shape space for which one body, say the back half, sweeps out a circular cone in the frame of the other body. Set $\mathbf{x} = g_f^{-1} g_b \mathbf{e}_3$. Then $\mathbf{x}$ sweeps out a geometric circle on the surface of the sphere. The north pole, $\mathbf{x} = (\mathbf{0}, \mathbf{0}, \mathbf{1})$ corresponds to the fully folded states. They denote the opening angle of the cone by β and the angle its axis makes with the north pole by α. One calculates that their cone is parameterized as $\mathbf{x}(t) = (sin(\beta)sin(t), sin(\beta)cos(\alpha)cos(t) +$

$cos(\beta)sin(\alpha), -sin(\beta)sin(\alpha)cos(t) + cos(\beta)cos(\alpha))$. Kane and Scher's variables are related to ours as follows.

Their variables	Our variables
ψ	χ
θ	t, parameter value along curve
T	$cos(\psi + \pi) =$ z component of curve
J/I	α
$\frac{S}{\sqrt{2}}$	$sin^2(\psi)\frac{d\theta}{dt}$

Trig identities imply that their $1 + T$ is our $2s^2$ and their $1 - T$ is our $2c^2$. Applying our parallel translation law ($d\chi = -\Gamma(\psi)d\theta$) to their curve yields, after some algebra, their differential equation for reorientation, equations (5) and (6) of [18].

6.5. The Maximally Symmetric Case. Here $\alpha = 1$ so that both bodies have a spherical inertial ellipsoid and

$$\Gamma = sin(\psi/2)d\theta.$$

Set $\phi = \psi/2$ and interpret (ϕ, θ) as the standard coordinates for the upper hemisphere of the unit sphere. Then Γ is the connection form for the Levi-Civita connection on the round sphere. Globally, it is the unique $SO(3)$-invariant connection for our principal $O(2)$-bundle $Q_{nt} \to S_{nt} = SO(3) \to \mathbb{R}P^2$.

The kinetic energy induced metric on shape space coincides (up to scale) with the unique $SO(3)$ invariant metric on $\mathbb{R}P^2$. In coordinates this is

$$ds^2 = d(\frac{\psi}{2})^2 + cos(\frac{\psi}{2})^2 d\theta^2.$$

Its corresponding normal Hamiltonian is

$$H_n = \frac{1}{2}(\frac{1}{a(\psi)}p_\psi^2 + \frac{1}{b(\psi)}(p_\theta - p_\chi\Gamma(\psi))^2$$

See §3. Upon lifting to the sphere we recognize this as the Hamiltonian which governs the motion of a particle of charge p_χ (a constant parameter) travelling on the round sphere with a monopole at its center. The solutions are "small circles" - curves of constant geodesic curvature- on S^2. These are the the cones of Kane and Scher which we just describe.

6.6. Can the Cat Flip for Free?. Consider the natural metric (kinetic energy restricted to zero angular momentum deformations) in the asymmetric case $\alpha \neq 1$. J. Marsden asked: can the cat "flip for free?" in this case. In other words are there any solutions to the free Hamiltonian (total kinetic energy) which are also extremals for the optimal control Hamiltonian (kinetic energy minus vertical kinetic energy)? If so, must they always involve collisions of the body halves?

The natural metric is calculated to be

$$ds^2 = I_1\{\frac{1+\beta c^2}{2}d\psi^2 + \frac{2\alpha c^2}{\alpha s^2 + c^2}d\theta^2\}$$

(Recall that the "s" on the right hand side means $sin(\frac{\psi}{2})$.) Consequently

$$H_n = \frac{1}{2}\{\frac{2}{1+\beta s^2}p_\psi^2 + \frac{\alpha s^2 + c^2}{2\alpha c^2}(p_\theta - \Gamma(\psi)p)^2\}$$

The class of free and controlled trajectories coincide precisely when $p_\chi = 0$. This is most easily seen by observing that p_χ represents the Lagrange multiplier needed to impose the velocity constraint "angular momentum equals zero". Alternatively, H_n = (horizontial Kinetic energy) = (total) - (vertical) kinetic energies, and the vertical term in linear in p_χ. (See [22].) When $p_\chi = 0$ we have $H_n = H$, the Hamiltonian for geodesic flow on $\mathbb{R}P^2$ with respect to the given metric ds^2. Since this metric is rotationally symmetric the usual analysis of geodesic flow on surfaces of revolution based on Clairut's theorem,

$$p_\theta = const.,$$

is valid. Call this constant μ and set

$$x = c^2 = cos(\frac{\psi}{2})$$

so that

$$\frac{\alpha s^2 + c^2}{4\alpha c^2} = \frac{1}{4x} + \frac{1-\alpha}{4\alpha} \equiv V(x)$$

V is a monotone decreasing function of x and since $0 \leq x \leq 1$ it is minimized when $x = 1$ in which case $\psi = 0$. These correspond to the collisions (the line at infinity). Now

$$H = g(x)p_\psi^2 + V(x)\mu^2$$

with $g(x)$ a positive function so that

$$V(x) \leq \frac{H}{\mu^2} = const.$$

with equality if and only if $p_\psi = 0$. (Look at the graph of V.) Note that $p_\psi = 0$ is equivalent to $\dot{\psi} = 0$. It follows that every free trajectory must pass through the the set where $x = 1$, i.e. the two body halves must collide. After this ψ increases until the above inequality is an equality. There x bounces back and again passes through the line at infinity , $x = 1$. In other words the 'radial' or ψ motion is always oscillatory and always passes through collision states $\psi = 0$. **In all free motions collisions must occur between the body halves!**

6.7. Impossibility of Rational or Elliptic Solutions with Natural Energy. Consider the case $\alpha \neq 1$, the case of non-maximal symmetry, with either the natural cost function used in the previous subsection, or the "maximally symmetric"cost $d\psi^2 + c^2 d\theta^2$. This is a completely integrable problem,as are all autonomous one degree of freedom systems.

The standard method of solution is to note that

$$\dot{\psi} = \frac{1}{a(\psi)} p_\psi$$

and then solve for $\dot{\psi}^2$ using the constants of the motion

$$\dot{\psi}^2 = \frac{2H_n}{a(\psi)} - \frac{b(\psi)}{a(\psi)}(p_\theta - \Gamma(\psi)p_\chi)^2$$

All terms are polynomial in sin $\frac{\psi}{2} = s$ so it is convenient to change variables to s, noting that

$$\dot{s}^2 = c^2\dot{\psi}^2 = (1-s^2)\dot{\psi}^2.$$

Then the equation reads

$$\dot{s}^2 = (1-s^2)\frac{2H_n}{a_1(s)} + \frac{b_1(s)}{a_1(s)}(p_\theta - \Gamma_1(s)p_\chi)^2$$

where $a_1(s) = a(\psi)$ etc...

Adding the fraction on the right hand side, one finds that the numerator has degree at least 6 in s, for generic constants of the motion (H_n, p_θ, p_χ). **Consequently the system is not integrable either in terms of elliptic or elementary functions. Hyperelliptic functions are needed.**

In the next and final section we show how a judicious yet simple choice of cost leads to a problem whose solution is immediate and elementary.

7. A Change of Variables to Reduce to the case of Maximal Symmetry

Reconsider the formula

$$\Gamma(\psi) = \frac{\alpha s}{c^2 + \alpha^2}$$

for the connection form **as a change of variables**

$$s = sin(\frac{\psi}{2}) \longmapsto \Gamma = \Gamma(s)$$

and define an angle $\phi, 0 \leq \phi \leq \pi$ by

$$\Gamma = sin(\phi/2)$$

Observe that $\Gamma(0) = 0, \Gamma(1) = 1$ and that the derivative $\frac{d\Gamma}{ds}$ is positive for $s \neq 0, 1$. It follows that this change of variables is an **invertible** orientation preserving map of the unit interval. In fact its inverse can be calculated by solving a quadratic equation. We get

$$s = \frac{-\alpha + \sqrt{\alpha^2 + 4\Gamma^2(1-\alpha)}}{2\Gamma(1-\alpha)}$$

(The choice of + sign in front of the square root coming from the quadratic formula is determined by checking the value of the function at $\Gamma = 0$. The leading term in the MacLaurin expansion is $\frac{1}{\alpha}\Gamma$.) Thus this change of variables also defines a diffeomorphism

$$\phi = \phi(\psi)$$

of the angle interval $[0, \pi]$ and thus a diffeomorphism Φ_α of $Q_{nt} = SO(3)$ given in coordinates by $(\chi, \psi, \theta) \mapsto (\chi, \phi(\psi), \theta)$.

By construction, under this change of variables

$$\Phi_\alpha^*(d\chi - \Gamma(\psi)d\theta) = d\chi - sin(\frac{\phi}{2})d\theta$$

which is the formula for the maximally symmetric connection under S0(3). Now define the cost function to be

$$\phi_\alpha^* ds^2$$

where $ds^2 = d\psi^2 + cos(\frac{\psi}{2})^2 d\theta^2$ is the maximally symmetric metric. In terms of the new variables this cost function is $d\phi^2 + cos(\frac{\phi}{2})^2 d\theta^2$. We have already found the extremals for the maximally symmetric case: they are circles on $\mathbb{R}P^2$! By general principles the extremals with respect to our cost function are the inverse image of these extremals under the map ϕ_α. The entire procedure is constructive.

Remark 7 (Credits). *The idea just presented had its germ in discussions with Sastry and company. They told me that what is really important for them is gett ing there, not getting there optimally. In other words, the control system is much more important than the cost function. Thus one should look for the cost function leading to the "simplest" optimal controls. The idea that the maximally symmetric cat's kinetic cost function might be such a function is due to Enos. The change of variables is mine.*

An algorithm for solving the steering problem is then to transform the initial and final points to the symmetric case using Φ_α, solve the problem in this case using the usual two step procedure, and then transform the resulting curve back using Φ_α. We recall the "usual two step procedure". Suppose for simplicity that the final shape x_1 and initial shape x_0 lie in the same affine chart and that with respect to the trivialization over this chart the final and initial elements can be connected, i.e. the holonomy is not in the disconnected part of the structure group $O(2)$. Thus we can work over the affine plane and take the structure group to be the circle subgroup generated by e_3. Let χ_1 and χ_0 be the final and initial orientations of the configurations relative to our coordinates. Then the two step procedure is :

Algorithm 1. *1. Get from the initial shape to the desired shape x_1 by travelling along a straight line in the affine plane (A geodesic on $\mathbb{R}P^2$ or any other standard choice of curve will work as well.) Calculate the new orientation (parallel transport) χ_2 at x_1 resulting from having travelled this line.*

2. Travel around the geometric circle whose oriented area with respect to the form $d(\sin(\psi/2)d\theta)$ is $\chi_2 - \chi_1$.

8. SUMMARY

The subject has come full circle. We took a long excursion from Kane and Scher's original model through bundles and gauge fields. In the end we found that the original solutions of Kane and Scher, after a judicious change of variables, are in certain senses both the optimal and the simplest solutions.

References

[1] R. Abraham and J. E. Marsden, [1978], **Foundations of Mechanics**, 2nd ed. Benjamin/Cummings.

[2] W. Ambrose and I.M. Singer [1953], *A theorem on holonomy*, Trans. AMS,75, 428-453

[3] V.I. Arnold, V.V. Kozlov, and A.I. Neishtatdt, [1988], **Dynamical Systems III**, v.3 in the Springer translation series Encyclopaedia of Mathematical Sciences, Springer- Verlag,.

[4] [1992] A. Bloch, M. Reyhanoglu, and N. McClamroch, *Control and Stabilization of Nonholonomic Dynamic Systems*, preprint, submitted to SIAM J. of Control and Optimization.

[5] R.W. Brockett, [1983] , *Nonlinear Control Theory and Differential Geometry*, Proc of the Int Congress of Mathematicians Warszawa.

[6] R.W. Brockett, [1981], *Control Theory and Singular Riemannian Geometry*, in New Directions in Applied Mathematics, P.J. Hilton and G.S.,Young, editors., Springer.

[7] C. Carathe"odory, [1967], **Calculus of Variations and Partial Differential Equations of the First Order**, v. 2. Holden-Day, S.F., CA.

[8] W.L. Chow, [1939], *Uber Systeme van Linearen partiellen Differentialgleichungen erster Ordnung*, Math. Ann, 117, 98-105.

[9] M. Enos, [1991] *Stationary Motions of the Time-Optimal Problem for a Force-free System of Two Symmetric Rigid Bodies in Three Space* , preprint, submitted to SIAM Journal of Control and Optimization.

[10] M. Enos, [1991] *Controllability of a System of Two Symmetric Rigid bodies in Three Space*, preprint, submitted to SIAM Journal of Control and Optimization.

[11] L.E. Faibusovich, [1988], *Explicitly Solvable nonlinear optimal controls*, Int'l J. Control, 48, **6**, 2507-26.

[12] A. Guichardet, [1984] *On rotation and vibration motions of molecules* Ann. Inst. H. Poincare, Phys. Theor., **40** no.3 329-342 .

[13] U. Hamensta dt [1986], *Sur Theorie von Carnot-Caratheodory-Metriken und ihren Anwendungen*, Doktorarbeit, Bonn.

[14] U. Hamensta,dt [1991], *Some Regularity Theorems for Carnot-Caratheodory Metrics* , J.D.G..

[15] R. Hermannn [1962], *Some Differential Geometric Aspects of the Lagrange Variational Problem*, Indiana Math. J., 634-673.

[16] R. Hermannn [1973], *Geodesics of Singular Riemannian metrics*, Bull. AMS 79 **4**, 780-782.

[17] L. Hsu, *Calculus of Variations via the Griffiths Formalism*, PhD thesis, Duke Univ., 1991.

[18] T. R. Kane and M. P. Scher, [1969], *A Dynamical Explanation of the Falling Cat Phenomenon*, Int'l J. Solids and Structures, 5, 663-670.

[19] S. Kobayashi and K. Nomizu [1963], **Foundations of Differential Geometry**, vol. 1. Interscience pub.

[20] J. Milnor, [1963], **Morse Theory**, Princeton Univ. Press.

[21] R. Montgomery [1992], *Gauge Theory and Optimal Control* , in **Nonholonomic Control**, conference proceedings of the IEEE Robotics Conference, Sacramento CA, 1991, Li and Canny ed.

[22] R. Montgomery, [1990], *Isoholonomic Problems and some applications*, Comm. Math. Phys. , vol. 128, 565-592.

[23] R. Montgomery, [1992], *Abnormal Optimal Controls and Some Open Problems in Nonholonomic Steering*, invited paper, for conference titled **Nonlinear Control Systems Design Symposium**, Bordeaux , France, June 24-26 1992, proceedings to appear in a Pergamom Press book.

[24] R. Montgomery, [1991], *Geodesics which do not satisfy the geodesic equations*, submitted to Math. Zeitschrift.

[25] G.J. Rademaker and J.W. G. ter Braak, [1935], *Das Umdrehen der fallenden Katze in der Luft*, Acta oto-lar. **23**, 313-343.

[26] S. Shankar Sastry and R. Montgomery, [1992], *The structure of optimal controls for a steering problem*, invited paper, for conference titled **Nonlinear Control Systems Design Symposium**, Bordeaux , France, June 24-26 1992, proceedings to appear in a Pergamom Press book.

[27] A. Shapere and F. Wilczek, [1987], *Self-propulsion at Low Reynolds number*, Phys. Rev. Lett. 58, 2051-54.

[28] A. Shapere, [1989], *Gauge Mechanics of Deformable Bodies*, PhD. thesis, Physics, Princeton.

[29] R. Strichartz [1983], *Sub-Riemannian Geometry*, J. Diff. Geom. 24, 221-263.

[30] F. Wilczek, [1988] *Gauge Theory of Deformable bodies*, Inst. for Adv. Studies, preprint -88/41.

[31] S.K. Wong , [1970], *Field and Particle Equations for the Classical Yang-Mills Field and Particles with Isotopic Spin*, Nuovo Cimento, 65a, 689-693.

[32] L.C. Young, **Lectures on the Calculus of Variations and Optimal Control Theory**, Chelsea, 1980.