SUSLOV PROBLEM WITH THE KLEBSH-TISSERAND POTENTIAL

SHENGDA HU, MANUELE SANTOPRETE

Abstract. In this paper, we study a nonholonomic mechanical system, namely the Suslov problem with the Klebsh-Tisserand potential. We analyze the topology of the level sets defined by the integrals in two ways: using an explicit construction and as a consequence of the Poincaré-Hopf theorem. We describe the flow on such manifolds.

1. Introduction

A Hamiltonian system on a $2n$-dimensional symplectic manifold is called completely integrable if it admits $n$ independent integrals of motion in involution. For such systems, if the common level sets $S_k$ of the integrals are compact, by the Liouville-Arnold theorem, the $S_k$ are invariant tori of dimension $n$, and the flow on the tori is isomorphic to a linear flow. The system is super-integrable if there are more than $n$ independent integrals of motions, and the invariant tori are of dimension less than $n$.

In the present paper, we are concerned with a family of dynamical systems, the so-called Suslov’s problem, that are not Hamiltonian, but exhibit important features of integrable and super-integrable Hamiltonian systems. As first formulated in [6], it describes the dynamics of a rigid body with a fixed point immersed in a potential field and subject to a nonholonomic constraint that forces the angular velocity component along a given direction in the body to vanish. Our analysis shows that such systems have invariant tori carrying linear flows, as well as other types of invariant submanifolds carrying generically periodic flows.

The topology of invariant submanifolds of this problem have been studied by Tatartnov [7, 8] using surgery methods, and Fernández-Bloch-Zenkov [2] using a generalization of the Poincaré-Hopf theorem to manifolds without boundary together with some detailed information about the geometry of the problem. It was shown that the invariant submanifolds of this problem can be surfaces of genus between zero and five.

We will provide two further approaches for understanding the topology of the submanifolds, as well as the flows. The first is a direct construction that uses a Morse theoretic reasoning and in our opinion provides a better understanding of the geometry of the problem than the other approaches. The second is an application of the classical Poincaré-Hopf theorem for manifolds without boundary and requires only knowledge of the number of connected components of the manifold. Furthermore, we give a detailed analysis of the flow on the invariant submanifolds and find that, for certain values of the parameters, the system admits an additional integral of motion. The information thus obtained leads to concrete understanding of the physical motion of the problem.

Suslov’s system is an example of an important class of nonholonomic systems, namely, the quasi-Chaplygin systems introduced in [1]. This system is Hamiltonizable in a very
shown in [1] that Suslov’s problem is a quasi-Chaplygin system with respect to a Lie group $G$, with a principal connection, and the distribution $\mathcal{D}$ is the horizontal bundle of the connection. Therefore, given a vector $Y \in T_{x}Q$, we have the decomposition $Y = Y_{h} + Y_{v}$, with $Y_{h} \in \mathcal{D}_{x}$ and $X_{v} \in \mathcal{V}_{x}$, where $\mathcal{V}$ is the vertical bundle, and $\mathcal{V}_{x}$ is the tangent space to the fiber at $x$. In addition one must require that the Lagrangian $L$ is $G$-invariant.

A Chaplygin system is a non-holonomic system $(Q, \mathcal{D}, L)$ where $Q$ has a principal bundle structure $Q \rightarrow Q/G$ with respect to a Lie group $G$, with a principal connection, and the distribution $\mathcal{D}$ is the horizontal bundle of the connection. Therefore, given a vector $Y \in T_{x}Q$, we have the decomposition $Y = Y_{h} + Y_{v}$, with $Y_{h} \in \mathcal{D}_{x}$ and $X_{v} \in \mathcal{V}_{x}$, where $\mathcal{V}$ is the vertical bundle, and $\mathcal{V}_{x}$ is the tangent space to the fiber at $x$. In addition one must require that the Lagrangian $L$ is $G$-invariant.

For Chaplygin systems the constrained Lagrangian $L_{c}(x, \dot{x}) = L(x, \dot{x})$ induces a Lagrangian $l : TM \rightarrow \mathbb{R}$ via the identification $TM \approx \mathcal{D}/G$ (note that $\mathcal{D}/G$ has the structure of a vector bundle with base space $Q/G$ and fiber $\mathbb{R}^{k}$, with $k = \dim(Q/G)$). Suppose the Legendre transform exists for $l$, which in local coordinates $q$ on $M$ is given by $(q, \dot{q}) \rightarrow \left(q, \frac{dt}{d\dot{q}}\right)$. Under the Legendre transform, the system of equations (1.1) gives rise to a first order dynamical system on $T^{*}M$ with corresponding vector field $X_{nh} = \{\cdot, H\}_{AP}$ for some function $H : T^{*}M \rightarrow \mathbb{R}$, where $\{\cdot, \cdot\}_{AP}$ is an almost Poisson bracket (i.e. a skew-symmetric bilinear operation on functions that satisfies the Leibniz identity but fails to satisfy the Jacobi identity). We say that the system is Chaplygin Hamiltonizable, if there is a nonvanishing function $f : Q/G \rightarrow \mathbb{R}$ such that

$$X_{nh} = f(q)X_{H}$$

with $X_{H} = \{\cdot, H\}$, and $\{\cdot, \cdot\} = \frac{1}{f}\{\cdot, \cdot\}_{AP}$ is a Poisson bracket.

For a quasi-Chaplygin system, if there is a function $f : Q/G \rightarrow \mathbb{R}$, nonvanishing in $Q \setminus S$, and such that $X_{nh} = f(q)X_{H}$ for some $H : T^{*}M \rightarrow \mathbb{R}$, we call the system quasi-Chaplygin Hamiltonizable. The Suslov problem considered in this article is quasi-Chaplygin Hamiltonizable with $f = \gamma_{3}$ ([1, 2]). For this type of systems, since the multiplier $f$ has zeroes, one hypothesis of Theorem 1 in [4] fails and thus the topology of invariant manifolds may differ from tori. Here we give an explicit description of the invariant manifolds. We expect that the more explicit approach taken in this article may be able to shed light on more general quasi-Chaplygin Hamiltonizable systems.
The paper is organized as follows. In Section 2 we present an overview of the Suslov’s problem in the Klebsh-Tisserand case. We give an elementary derivation of the equations of motion (see [1] for a derivation based on a Lagrangian approach). In Section 3 we give an explicit construction that allows us to determine the topology of the level surfaces $S_k$. Then, in Section 4 we study the flow of the system on the surfaces $S_k$, and we find that, for some parameter values there is one additional integral of motion. We then find and classify the critical points of the vector field. We use the Poincaré-Hopf theorem to give an alternative way to determine the topology of the surfaces. In Section 5 we use the topology of $S_k$ and the results on the dynamics of the problem to describe how the rigid body moves in the three dimensional physical space.

2. Suslov’s Problem with a Klebsh-Tisserand Potential

The Suslov problem describes the motion of a rigid body with a fixed point subject to a nonholonomic constraint. Wagner [9] suggested the following implementation of Suslov’s model. He considered a rigid body, with a fixed point $O$, moving inside a spherical shell. The rigid body is attached at $O$ with a spherical hinge so that it can turn around this point. The nonholonomic constraint is realized by considering two rigid caster wheels attached to the rigid body by a rod (see figure 1). These wheels force the angular velocity component along a direction orthogonal to the rod to vanish.

![Figure 1. An implementation of the Suslov problem suggested by Wagner.](image)

In this section we give an elementary derivation of the equations of motion for the Suslov’s problem. These equations can also be obtained from a Lagrangian [1]. We begin by discussing the Euler equations for a rigid body without constraint. Denote
by \(\{e_1, e_2, e_3\}\) a right handed orthonormal basis of \(\mathbb{R}^3\), called the spatial frame. The coordinates of a point \(P\) is the spatial frame is denoted \(x\), which is also called the spatial vector of \(P\). Let \(\{E_1, E_2, E_3\}\) be a right handed orthonormal frame, the body frame, defined by the three principal axis and let \(X\) denote the body vector of the point \(P\) in this frame. We have

\[
x = RX,
\]

where \(R\) is a rotation matrix.

Let \(\omega\), \(\pi\) and \(\tau\) be the spatial vector of angular velocity, angular momentum, and torque, respectively. The corresponding body vectors are then given by

\[
\Omega = R^{-1} \omega, \quad \Pi = R^{-1} \pi \quad \text{and} \quad T = R^{-1} \tau
\]

Note that if \(I = \text{diag}(I_1, I_2, I_3)\) denotes the body inertia tensor then we can also write \(\Pi = I \Omega\). Now \(\tau = \dot{\pi}\), the second cardinal equation of dynamics, can be rewritten as

\[
RT = \tau = \frac{d}{dt} (R \Pi) = R \dot{\Pi} + R \dot{R} \Pi \Rightarrow \dot{\Pi} = -R^{-1} R \dot{R} \Pi + T
\]

Let \(\hat{\Omega} = R^{-1} \dot{R}\) be the hat map, then we have

\[
\hat{\Omega} \Pi = \Omega \times \Pi
\]

which gives

\[
\dot{\Pi} = \Pi \times \Omega + T
\]

Let \(\alpha\), \(\beta\), and \(\gamma\) be the body vectors of \(e_1\), \(e_2\), and \(e_3\) respectively, e.g.

\[
\gamma = R^{-1} e_3
\]

Then we obviously have

\[
\|\gamma\|^2 = \gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1.
\]

Suppose the rigid body is placed in a force field with potential energy

\[
u(X) = u \left( \langle X, \gamma \rangle \right),
\]

where \(\langle ,\rangle\) is the Euclidean inner product. Then the total potential energy of the rigid body is

\[
U(\gamma) = \int_B u \left( \langle X, \gamma \rangle \right) d^3X
\]

and the total body force and torque are

\[
F = -\int_B \frac{\partial u}{\partial X} d^3X, \quad \text{and} \quad T = -\int_B \frac{\partial u}{\partial X} \times X d^3X,
\]

respectively. We have

\[
\frac{\partial u}{\partial X} \times X = \frac{\partial u}{\partial \langle X, \gamma \rangle} \frac{\partial \langle X, \gamma \rangle}{\partial X} \times X = \left( \frac{\partial u}{\partial \langle X, \gamma \rangle} X \right) \times \gamma,
\]

and similarly

\[
\frac{\partial u}{\partial \gamma} = \frac{\partial u}{\partial \langle X, \gamma \rangle} \frac{\partial \langle X, \gamma \rangle}{\partial \gamma} = \frac{\partial u}{\partial \langle X, \gamma \rangle} X
\]

which gives

\[
\frac{\partial u}{\partial X} \times X = -\frac{\partial u}{\partial \gamma} \times \gamma.
\]
Since $\gamma$ does not depend on $X$, integrating yields the following expression for the torque:

$$ T = \frac{\partial U}{\partial \gamma} \times \gamma $$

and thus the dynamics equations are

$$ \dot{\Pi} = \Pi \times \Omega + \frac{\partial U}{\partial \gamma} \times \gamma, \quad \dot{\gamma} = \gamma \times \Omega $$

Let $A$ be a fixed unit body vector, and consider Suslov’s nonholonomic constraint

$$ \langle \Omega, A \rangle = 0 $$

Subjecting the rigid body to this constraint is equivalent to adding a torque $\lambda A$. The rigid body subject to this torque moves according to

$$ \dot{\Pi} = \Pi \times \Omega + \frac{\partial U}{\partial \gamma} \times \gamma + \lambda A, \quad \dot{\gamma} = \gamma \times \Omega, \quad \langle \Omega, A \rangle = 0, $$

where the first equation can also be written as

\begin{equation}
\dot{\Omega} = \Omega \times \Omega + \frac{\partial U}{\partial \gamma} \times \gamma + \lambda A.
\end{equation}

Differentiating the constraint gives $\langle A, \dot{\Omega} \rangle = 0$, and substituting (2.1) into the constraint equation yields

$$ \langle A, \Pi^{-1} \left( \Omega \times \Omega + \frac{\partial U}{\partial \gamma} \times \gamma + \lambda A \right) \rangle = 0 $$

and hence

$$ \lambda = -\frac{\langle A, \Pi^{-1} \left( \Omega \times \Omega + \frac{\partial U}{\partial \gamma} \times \gamma \right) \rangle}{\langle A, \Pi^{-1} A \rangle}. $$

Let us consider the case $A = E_3$, so that the constraint is simply $\Omega_3 = 0$. Then the equation of the rigid body subject to Suslov’s nonholonomic constraint are

$$ I_1 \dot{\Omega}_1 = \gamma_2 \frac{\partial U}{\partial \gamma_3} - \gamma_3 \frac{\partial U}{\partial \gamma_2}, \quad I_2 \dot{\Omega}_2 = \gamma_3 \frac{\partial U}{\partial \gamma_1} - \gamma_1 \frac{\partial U}{\partial \gamma_3}, $$

and

$$ \dot{\gamma}_1 = -\gamma_3 \Omega_2, \quad \dot{\gamma}_2 = \gamma_3 \Omega_1, \quad \dot{\gamma}_3 = \gamma_1 \Omega_2 - \gamma_2 \Omega_1, $$

where we omitted the equation for $\dot{\Omega}_3$ since it is used only to determine $\lambda$ and can be omitted. We consider the Klebsh-Tisserand case of the Suslov problem, i.e.

$$ U(\gamma) = \frac{1}{2} \left( B_1 \gamma_1^2 + B_2 \gamma_2^2 \right) $$
Then the equations of motion in terms of the momenta are given by

\[
\begin{align*}
\dot{\Pi}_1 &= -B_2 \gamma_2 \gamma_3 \\
\dot{\Pi}_2 &= B_1 \gamma_1 \gamma_3 \\
\dot{\gamma}_1 &= -\gamma_3 \frac{\Pi_2}{I_2} \\
\dot{\gamma}_2 &= \gamma_3 \frac{\Pi_1}{I_2} \\
\dot{\gamma}_3 &= \gamma_1 \frac{\Pi_2}{I_2} - \gamma_2 \frac{\Pi_1}{I_2}.
\end{align*}
\]

The following functions are easily seen to be integrals of motion of the equations above

\[
F_1 := \frac{\Pi_1^2}{I_1} + B_2 \gamma_2^2, \quad F_2 := \frac{\Pi_2^2}{I_2} + B_1 \gamma_1^2.
\]

Understanding of the Suslov problem now reduces to understanding of the flows on the level surfaces \(F_1^{-1}(K_1) \cap F_2^{-1}(K_2)\) defined by the integrals of motion. It is convenient to perform the following change of variables:

\[
m_1 = -\frac{\Pi_2}{I_2}, \quad b_1 = \frac{B_1}{I_2}, \quad k_1 = \frac{K_1}{I_2}; \quad m_2 = -\frac{\Pi_1}{I_1}, \quad b_2 = \frac{B_2}{I_1}, \quad k_2 = \frac{K_2}{I_1}.
\]

and consider the system as defined in \(\mathbb{R}^5\), with coordinates \((m_1, m_2, \gamma_1, \gamma_2, \gamma_3)\), subject to the restriction

\[
\gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1
\]

Then the integrals of motion become \(f_1 = m_1^2 + b_1 \gamma_1^2\) and \(f_2 = m_2^2 + b_2 \gamma_2^2\). We write down the equations defining the level surfaces \(S_k := f_1^{-1}(k_1) \cap f_2^{-1}(k_2)\):

\[
\begin{align*}
\begin{cases}
m_1^2 + b_1 \gamma_1^2 = k_1 \\
m_2^2 + b_2 \gamma_2^2 = k_2 \\
\gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1
\end{cases}
\end{align*}
\]

We also write the equations of motion in the new coordinates:

\[
\begin{align*}
\dot{m}_1 &= -b_1 \gamma_1 \gamma_3 \\
\dot{m}_2 &= b_2 \gamma_2 \gamma_3 \\
\dot{\gamma}_1 &= m_1 \gamma_3 \\
\dot{\gamma}_2 &= -m_2 \gamma_3 \\
\dot{\gamma}_3 &= \gamma_2 m_2 - \gamma_1 m_1.
\end{align*}
\]

We denote by \(X = (X_1, X_2, X_3, X_4, X_5) : \mathbb{R}^5 \rightarrow \mathbb{R}^5\), the vector field associated with the equations above.
3. Topology of the level surfaces \( S_k \) via an explicit construction

In this section we want to describe the topology of the level surfaces \( S_k \) by giving an explicit construction. Our method differs from the surgery approach pioneered by Tatarinov \([7, 8]\). Note that while it is difficult to find the original work of Tatarinov, an exposition of his method can be found in \([3]\) and \([2]\).

We start by finding the values of \( k = (k_1, k_2) \) for which \( S_k \) is non-singular.

**Lemma 3.1.** The subspace \( S_k \) is a smooth manifold of dimension 2 iff \( k_1 k_2 \neq 0 \) and all of the following holds: \( k_1 \neq b_1, k_2 \neq b_2 \) and \( \frac{k_1}{b_1} + \frac{k_2}{b_2} \neq 1 \).

**Proof.** The matrix formed by the gradients of the defining equations is

\[
\begin{pmatrix}
m_1 & 0 & b_1 \gamma_1 & 0 & 0 \\
0 & m_2 & 0 & b_2 \gamma_2 & 0 \\
0 & 0 & \gamma_1 & \gamma_2 & \gamma_3 \\
\end{pmatrix}
\]

and \( S_k \) is smooth iff matrix above to have maximal rank 3 at all points on \( S_k \). Obviously, the matrix is of full rank when \( m_1 m_2 \neq 0 \); while it is degenerate when \( k_1 k_2 = 0 \).

For \( m_1 = m_2 = 0 \), we have, \( b_1 \gamma_1^2 = k_1 \) and \( b_2 \gamma_2^2 = k_2 \). It follows that

\[
\gamma_3^2 = 1 - \frac{k_1}{b_1} - \frac{k_2}{b_2}
\]

The rank 3 condition requires \( \frac{k_1}{b_1} + \frac{k_2}{b_2} \neq 1 \).

For \( m_1 = 0 \) but \( m_2 \neq 0 \), we have \( b_1 \gamma_1^2 = k_1 \). Rank 3 condition requires one of \( \gamma_2 \) or \( \gamma_3 \) is non-zero. We have \( \gamma_2 = \gamma_3 = 0 \Longleftrightarrow b_1 = k_1 \), thus full rank implies \( k_1 \neq b_1 \).

The case \( m_1 \neq 0 \) but \( m_2 = 0 \) gives \( k_2 \neq b_2 \). \( \square \)

A simple consequence of the Lemma is that the topology of the level sets \( S_k \) can only change at values of \( k \) where \( S_k \) is singular. This is made precise in the following corollary.

**Corollary 3.2.** The first quadrant in the \((k_1, k_2)\)-plane is divided into 5 regions by

\[
k_1 = b_1, k_2 = b_2 \text{ and } \frac{k_1}{b_1} + \frac{k_2}{b_2} = 1
\]

The subspace \( S_k \) has the same topological type for \( k \) in each region (c.f. Figure 2).

The level surfaces \( S_k \) are complete intersections. In \((2.2)\), the first two equations define a 2-torus \( T^2_k \subset \mathbb{R}^4 \), with coordinates \((m_1, m_2, \gamma_1, \gamma_2)\), while the last equation defines the unit 2-sphere \( S^2 \subset \mathbb{R}^3 \), with coordinates \((\gamma_1, \gamma_2, \gamma_3)\). It is therefore natural to study the level sets \( S_k \) by analyzing their projections onto these well understood surfaces. The projection onto the torus will be studied in the next subsection and it will be crucial in determining the topology of the level surfaces. The projection onto the unit 2-sphere \( S^2 \) will be studied in Subsection 5.1 to explain the connection between the topology of the \( S_k \) and the motion of the rigid body in physical space.
3.1. Projection onto the torus. In order to describe the projection of the level surfaces $S_k$ onto the 2-torus $T^2_k$ it is convenient to use a standard parametrization and describe the torus as the square flat torus. Since the first equation in (2.2) defines an ellipse in the $(m_1, \gamma_1)$-plane and the second equation defines an ellipse in the $(m_2, \gamma_2)$-plane, we parametrize these ellipses by the angle in the respective polar coordinates:

$$
\begin{align*}
    m_1 &= \sqrt{k_1} \cos \theta_1 \\
    \gamma_1 &= \sqrt{\frac{k_1}{b_1}} \sin \theta_1 \\
    m_2 &= \sqrt{k_2} \sin \theta_2 \\
    \gamma_2 &= \sqrt{\frac{k_2}{b_2}} \cos \theta_2,
\end{align*}
$$

Here the square $[-\frac{\pi}{2}, \frac{3\pi}{2}] \times [0, 2\pi]$ is viewed as the square flat torus $T^2$ by identifying the top side of the square with the bottom side, and the left side with the right side.

Then the parametrization above defines the following isomorphism from $T^2_k$ to the standard torus $T^2$:

$$
\varphi_k : T^2_k \cong T^2 : (m_1, m_2, \gamma_1, \gamma_2) \mapsto (\theta_1, \theta_2)
$$

By definition, for each $k$, we see that $S_k \subset T^2_k \times \mathbb{R}$, where the coordinate on the second factor is given by $\gamma_3$. Let $p_k : S_k \rightarrow T^2_k$ be the projection induced by the projection of $T^2_k \times \mathbb{R}$ to the first factor. The dependence of $S_k$ on $k$ can be described using $\varphi_k \circ p_k$.

To describe the projection of the surfaces $S_k$ onto the torus (or more precisely the image of $S_k$ under the map $\phi_k \circ p_k$) it is convenient to introduce the function $g_k : T^2 \rightarrow \mathbb{R}$ defined as

$$
g_k(\theta_1, \theta_2) = \frac{k_1}{b_1} \cos^2 \theta_1 + \frac{k_2}{b_2} \sin^2 \theta_2
$$

and let $k = (k_1, k_2)$. We denote by $U_k$ the subset of $T^2$ consisting of all points at which $g_k$ takes values greater than the real number $\epsilon_k$, that is we set

$$
U_k = \left\{ (\theta_1, \theta_2) \in T^2 : g_k > \epsilon_k = \frac{k_1}{b_1} + \frac{k_2}{b_2} - 1 \right\} \subseteq T^2
$$

and let $\partial U_k = g_k^{-1}(\epsilon_k)$. Let $\overline{U}_k = U_k \cup \partial U_k$ be the closure of $U_k$. Figure 3 shows the set $\overline{U}_k$ for various values of $k$.

The following Lemma shows that the set $\overline{U}_k$ is the image of $S_k$ under the map $\varphi_k \circ p_k$ and gives a characterization of such image.
Lemma 3.3. For all $k$, we have $\mathcal{U}_k = \varphi_k \circ p_k(S_k)$. The map $\varphi_k \circ p_k : S_k \to \mathcal{U}_k$ is 2-to-1 over the interior $U_k$ and is 1-to-1 over the boundary $\partial U_k$, if $\partial U_k \neq \emptyset$. Thus, $S_k$ is homomorphic to the surface obtained by attaching two copies of $\mathcal{U}_k$ along $\partial U_k$.

Proof. The equation defining $S_k$ in $T^2_k \times \mathbb{R}$ is $\gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1$. It follows that $\varphi_k \circ p_k(S_k)$ is the subset of $T^2$ defined by

$$1 \geq \gamma_1^2 + \gamma_2^2 = \frac{k_1}{b_1} \sin^2 \theta_1 + \frac{k_2}{b_2} \cos^2 \theta_2$$

This is exactly $\mathcal{U}_k$. Over $U_k$, the strict inequality above holds, which implies that $\gamma_3 \neq 0$ takes 2 distinct values. It follows that $\varphi_k \circ p_k$ is 2-to-1 over $U_k$. Over the boundary $\partial U_k$, the equality holds and it implies that $\gamma_3 = 0$. Thus $\varphi_k \circ p_k$ is 1-to-1 along $\partial U_k$ whenever it is not empty. \qed

A consequence of this result is that we can use the shape of the set $\mathcal{U}_k$ to characterize the geometry and the topology of the surfaces $S_k$. The following Lemma describes some feature of the function $g_k$ that can be used to describe the shape of the sets $\mathcal{U}_k$.

Lemma 3.4. The smooth function $g_k$ on $T^2$ is a Morse function. The 16 critical points are independent of $k$, with 4 critical points on each of the 4 critical levels:

1. Minimums at $\left\{-\frac{\pi}{2}, \frac{\pi}{2}\right\} \times \{0, \pi\}$, with $g_k = 0$. 

Figure 3. The set $\mathcal{U}_k$ for various values of $k$. 

(a) $\mathcal{U}_k$ for $k \in D_1$  
(b) $\mathcal{U}_k$ for $k \in D_2$  
(c) $\mathcal{U}_k$ for $k \in D_3$  
(d) $\mathcal{U}_k$ for $k \in D_4$  
(e) $\mathcal{U}_k$ for $k \in D_5$
(2) Saddles at \(-\frac{\pi}{2}, \frac{\pi}{2}\) \(\times\) \(\left\{\frac{\pi}{2}, \frac{3\pi}{2}\right\}\), with \(g_k = \frac{k_1}{b_1}\).

(3) Saddles at \(\{0, \pi\} \times \{0, \pi\}\), with \(g_k = \frac{k_2}{b_2}\).

(4) Maximums at \(\{0, \pi\} \times \left\{\frac{\pi}{2}, \frac{3\pi}{2}\right\}\), with \(g_k = \frac{k_1}{b_1} + \frac{k_2}{b_2}\).

Proof. The statement follows from straightforward computations. □

We can now use Lemmata 3.3 and 3.4 to completely characterize the surfaces \(S_k\).

Proposition 3.5. The topology of the surfaces \(S_k\) is described below:

- For \(k \in D_1\), \(S_k\) is isomorphic to two copies of \(T^2\).
- For \(k \in D_2\), \(S_k\) is a genus 5 surface.
- For \(k \in D_3\) or \(D_4\), \(S_k\) is isomorphic to two copies of \(T^2\).
- For \(k \in D_5\), \(S_k\) is isomorphic to four copies of \(S^2\).

Proof. The results follow from understanding the set \(U_k\) using Lemma 3.4. For \(k \in D_1\) we have \(\frac{k_1}{b_1} + \frac{k_2}{b_2} < 1\), which gives \(\varepsilon_k < 0\). Since \(g_k \geq 0\), we see that \(\partial U_k = \emptyset\), see Figure 3a. Thus \(S_k\) is isomorphic to two copies of \(T^2\).

For \(k \in D_2\), we have \(k_1 < b_1, k_2 < b_2, \frac{k_1}{b_1} + \frac{k_2}{b_2} > 1 \implies 0 < \varepsilon_k < \min\left\{\frac{k_1}{b_1}, \frac{k_2}{b_2}\right\}\). It follows by Lemma 3.4 that the set \(U_k\) is isomorphic to \(T^2 \setminus 4D^2\), and \(\partial U_k \cong 4S^1\), see Figure 3b. Lemma 3.3 implies that \(S_k\) is isomorphic to two copies of \(T^2\) connect sum at 4 distinct points, i.e. a genus 5 surface.

For \(k \in D_3\), we have \(k_1 > b_1, k_2 < b_2 \implies \frac{k_2}{b_2} < \varepsilon_k < \frac{k_1}{b_1}\).

Then Lemma 3.4 implies that \(U_k\) consists of two components \(C_{k,1} \cup C_{k,2}\), each of which is isomorphic to \(S^1 \times (0, 1)\), and \(\partial U_k \cong 4S^1\), see Figure 3c. Apply Lemma 3.3, we see that \(S_k\) has two components as well, each of which is isomorphic to a 2-torus. The argument for \(k \in D_4\) is similar, where we have \(k_1 < b_1, k_2 > b_2 \implies \frac{k_1}{b_1} < \varepsilon_k < \frac{k_2}{b_2}\)
and \(U_k\) again consists of two components, see Figure 3d. Again, \(S_k\) in this case has two components and each is isomorphic to a 2-torus.

For \(k \in D_5\), we have \(k_1 > b_1, k_2 > b_2 \implies \max\left\{\frac{k_1}{b_1}, \frac{k_2}{b_2}\right\} < \varepsilon_k < \frac{k_1}{b_1} + \frac{k_2}{b_2}\).

Lemma 3.4 implies that \(U_k\) consists of 4 components \(H_{k,j}\) for \(j = 1, 2, 3, 4\), each of which is isomorphic to \(D^2\), the 2-disk, and \(\partial U_k \cong 4S^1\), see Figure 3e. With Lemma 3.3, we find that \(S_k\) has four components, each of which is isomorphic to an \(S^2\). □
4. Dynamics on the level surfaces $S_k$

The projection to the torus as described in the previous section also provides us with detailed information on the Suslov flow.

4.1. Linear flow on tori. We consider the region $D_1$ in Figure 2, where the level sets are two tori. From the proof of Proposition 3.5 each component of the level set at $(k_1, k_2) \in D_1$ is diffeomorphic to the torus in $\mathbb{R}^4$ with coordinates $(m_1, m_2, \gamma_1, \gamma_2)$ given by

$$\{m_1^2 + b_1 \gamma_1^2 = k_1, m_2^2 + b_2 \gamma_2^2 = k_2\}$$

Each equation above defines an ellipse in the plane, which, as we have seen, can be parametrized by introducing polar coordinates in each plane (3.2). In these coordinates, on the level surface, the Suslov flow (2.3) takes the form

$$\dot{\theta}_1 = \sqrt{b_1} \gamma_3, \quad \dot{\theta}_2 = \sqrt{b_2} \gamma_3$$

Thus the Suslov flow projects to a linear flow with slope $\sqrt{b_2/b_1}$ on the torus, which is periodic when the ratio is a rational number. On the square flat torus $T^2$, the projected flow is given by pieces of straight lines with slope $\sqrt{b_2/b_1}$, see figure 4a.

![Figure 4](image)

(a) $b_1 = 4$, $b_2 = 1$, $k_1 = 1$, and $k_2 = 0.5$  
(b) $b_1 = 4$, $b_2 = 1$, $k_1 = 4.4$, and $k_2 = 1.1$

**Figure 4.** (a) The flow for $k \in D_1$. In this case have rational slope $\sqrt{b_2/b_1} = 1/2$, and all the orbits are periodic. The picture shows four periodic orbits. (b) The flow for $k \in D_3$. There are 8 critical points. All the other orbits are periodic.
4.2. **Additional integral of motion.** When $\sqrt{\frac{b_2}{b_1}} \in \mathbb{Q}$ the system admits another integral of motion, which implies as well that Suslov flow on $S_k$ is periodic for generic $k$ in this case. We describe it first for $b_1 = b_2 = b$.

**Proposition 4.1.** When $b_1 = b_2 = b$, the Suslov flow has the following as an integral of motion:

$$f_3 = m_1 m_2 - b \gamma_1 \gamma_2$$

**Proof.** Straightforward verification by taking derivative with respect to $t$. \qed

It’s readily verified that the level sets of $f_3$ define the periodic flow on the tori for $k \in D_1$ when $b_1 = b_2$. In general, the new integral of motion is a higher degree polynomial in $m_i$’s and $\gamma_i$’s.

**Proposition 4.2.** Suppose that the ratio $\sqrt{b_1} : \sqrt{b_2}$ is rational, then there is an integral of motion $f_3$, given by a polynomial of $(\gamma_1, \gamma_2, m_1, m_2)$.

**Proof.** For a given $k \in D_1$, rewrite the flow equations (2.3) in the $(\theta_1, \theta_2)$-coordinates:

\[
\begin{align*}
  m_1 &= \sqrt{k_1 \cos \theta_1} \\
  \gamma_1 &= \sqrt{k_1} \sin \theta_1 \\
  m_2 &= \sqrt{k_2 \sin \theta_2} \\
  \gamma_2 &= \sqrt{k_2} \cos \theta_2
\end{align*}
\]

where $-\frac{\pi}{2} \leq \theta_1 \leq \frac{3\pi}{2}$ and $0 \leq \theta_2 \leq 2\pi$, and we have

$$\dot{\theta}_1 = \sqrt{b_1} \gamma_3, \dot{\theta}_2 = \sqrt{b_2} \gamma_3$$

Let $p, q \in \mathbb{Z}$ be integers such that

$$\sqrt{\frac{b_1}{b_2}} = \frac{p}{q}$$

then we see that $q \theta_1 - p \theta_2$ is a constant along the flow

$$q \dot{\theta}_1 - p \dot{\theta}_2 = (q \sqrt{b_1} - p \sqrt{b_2}) \gamma_3 = 0$$

Furthermore, we can express trigonometric functions of $q \theta_1 - p \theta_2$ as a degree $p + q$ polynomial in $m_1, m_2, \gamma_1, \gamma_2$, involving also $k_1$ and $k_2$. For example, let $z_1 = e^{i\theta_1}$ and $z_2 = e^{i\theta_2}$, then

$$\cos(q \theta_1 - p \theta_2) = \mathbb{R}(z_1^q z_2^{-p}) = \mathbb{R}((\cos \theta_1 + i \sin \theta_1)^q (\cos \theta_2 - i \sin \theta_2)^p)$$

It follows that

$$f_3 = \mathbb{R} \left( (m_1 + i \sqrt{b_1} \gamma_1)^q \left( \sqrt{b_2} \gamma_2 - i m_2 \right)^p \right) = k_1^q k_2^p \cos(q \theta_1 - p \theta_2)$$

is a constant along the flows for $k \in D_1$. It is straightforward to verify by direct differentiation that $(m_1 + i \sqrt{b_1} \gamma_1)^q \left( \sqrt{b_2} \gamma_2 - i m_2 \right)^p$ is a constant along the flow independent of $k$. Thus $f_3$ is an integral of motion, which is a degree $p + q$ real polynomial in $(\gamma_1, \gamma_2, m_1, m_2)$. \qed
4.3. Critical Points. Critical points of the flow of the Suslov problem can be obtained by a simple geometric argument. We observe that the critical points are precisely where the level sets $\partial U_k$ are tangent to the linear flow. Thus, in $(\theta_1, \theta_2)$ coordinates, the critical points are exactly the solutions to the following system of equations:

\[
\begin{align*}
\frac{k_1}{b_1} \sin^2 \theta_1 + \frac{k_2}{b_2} \cos^2 \theta_2 &= 1 \\
k_1 b_2 \sin \theta_1 \cos \theta_1 = \sqrt{\frac{b_2}{b_1}} k_2 \sin \theta_2 \cos \theta_2
\end{align*}
\]

The second equation simplifies to

\[
\frac{k_1^2}{b_1} \sin^2 \theta_1 (1 - \sin^2 \theta_1) = \frac{k_2^2}{b_2} (1 - \cos^2 \theta_2) \cos^2 \theta_2
\]

Using the first equation and the fact that $\gamma_i^2 = \frac{k_1}{b_1} \sin^2 \theta_1$, we obtain the following quadratic equation in $\gamma_i^2$, which can be explicitly solved:

\[
(4.1) \quad (b_1 - b_2) \gamma_i^4 - (k_1 + k_2 - 2b_2) \gamma_i^2 + (k_2 - b_2) = 0.
\]

When $b_1 = b_2 := b$, (4.1) reduces to

\[
(k_1 + k_2 - 2b) \gamma_i^2 + (k_2 - b) = 0 \implies \gamma_i = \pm \gamma_i^*, \text{ where } \gamma_i^* = \sqrt{\frac{k_2 - b}{k_1 + k_2 - 2b}}
\]

Let

\[
\gamma_2^* = \sqrt{\frac{k_1 - b}{k_1 + k_2 - 2b}} \text{ and } t = \sqrt{k_1 + k_2 - b}
\]

then the critical points in this case are given in Table 1 below:

| Region | Value of Parameters | # of cp | critical points: $(m_1, m_2, \gamma_1, \gamma_2, \gamma_3)$ |
|--------|---------------------|---------|-----------------------------------------------------|
| $D_1$  | $k_1 + k_2 < b$     | 0       | $\pm (\pm t \gamma_2^*, \pm t \gamma_1^*, \pm \gamma_1^*, \pm \gamma_2^*, 0)$, $\pm (\pm t \gamma_2^*, \mp t \gamma_1^*, \mp \gamma_1^*, \pm \gamma_2^*, 0)$ |
| $D_2$  | $k_1 + k_2 > b$, $k_1 < b$, $k_2 < b$ | 8       | $\pm (\pm t \gamma_2^*, \pm t \gamma_1^*, \pm \gamma_1^*, \pm \gamma_2^*, 0)$, $\pm (\pm t \gamma_2^*, \mp t \gamma_1^*, \mp \gamma_1^*, \pm \gamma_2^*, 0)$ |
| $D_3$  | $k_1 > b$, $k_2 < b$ | 0       | $\pm (\pm t \gamma_2^*, \pm t \gamma_1^*, \pm \gamma_1^*, \pm \gamma_2^*, 0)$, $\pm (\pm t \gamma_2^*, \mp t \gamma_1^*, \mp \gamma_1^*, \pm \gamma_2^*, 0)$ |
| $D_4$  | $k_1 < b$, $k_2 > b$ | 0       | $\pm (\pm t \gamma_2^*, \pm t \gamma_1^*, \pm \gamma_1^*, \pm \gamma_2^*, 0)$, $\pm (\pm t \gamma_2^*, \mp t \gamma_1^*, \mp \gamma_1^*, \pm \gamma_2^*, 0)$ |
| $D_5$  | $k_1 > b$, $k_2 > b$ | 8       | $\pm (\pm t \gamma_2^*, \pm t \gamma_1^*, \pm \gamma_1^*, \pm \gamma_2^*, 0)$, $\pm (\pm t \gamma_2^*, \mp t \gamma_1^*, \mp \gamma_1^*, \pm \gamma_2^*, 0)$ |

**Table 1.** Critical points for $b_1 = b_2$. See Figure 2 for a definition of the regions $D_i$.

Suppose that $b_1 \neq b_2$. As a quadratic equation in $\gamma_1^2$, the discriminant of (4.1) is

\[
\Delta = (k_1 + k_2 - 2b_2)^2 - 4(b_1 - b_2)(k_2 - b_2)
\]

The solutions of equation (4.1) are $\gamma_i = \pm \Gamma_i^{-1}$, $\pm \Gamma_i^{+1}$ with

\[
\Gamma_i^{-1} = \sqrt{\frac{(k_1 + k_2 - 2b_2) - \sqrt{\Delta}}{2(b_1 - b_2)}}, \quad \Gamma_i^{+1} = \sqrt{\frac{(k_1 + k_2 - 2b_2) + \sqrt{\Delta}}{2(b_1 - b_2)}}
\]
which can be real or complex depending on the value of the parameters.

\[
\Gamma^\pm_2 = \sqrt{1 - (\Gamma^\pm_1)^2}, \quad \Lambda_i^\pm = \sqrt{k_i - b_i(\Gamma^\pm_1)^2} \quad \text{and} \quad \Lambda_2^\pm = \sqrt{k_2 - b_2(\Gamma^\pm_2)^2}
\]

We can then describe the critical points for \(b_1 > b_2\) in Table 2 below, where the regions are as labeled in Figure 5.

| Region | Value of Parameters | # | critical points: \((m_1, m_2, \gamma_1, \gamma_2, \gamma_3)\) |
|--------|---------------------|---|-----------------------------------------------|
| \(D_1\) | \(k_2 < b_2\) | \(\Delta > 0\) | \(k_1 > b_1\) | 0 | \((-1)^k \Lambda_1^+, \gamma_1 L, (-1)^j \Lambda_2^+, \gamma_2, \gamma_3, (-1)^j L\) |
| \(D_2\) | \(k_2 < b_2\) | \(\Delta > 0\) | \(k_1/b_1 + k_2/b_2 > 1, k_1 < b_1\) | 8 | \((-1)^k \Lambda_1^+, \gamma_1 L, (-1)^j \Lambda_2^+, \gamma_2, \gamma_3, (-1)^j L\) |
| \(D_3\) | \(k_2 < b_2\) | \(\Delta > 0\) | \(k_1/b_1 + k_2/b_2 < 1\) | 0 | \((-1)^k \Lambda_1^+, \gamma_1 L, (-1)^j \Lambda_2^+, \gamma_2, \gamma_3, (-1)^j L\) |
| \(D_4\) | \(k_2 > b_2\) | \(\Delta > 0\) | \(k_1 < b_1, k_1 + k_2 > 2b_1\) | 8 | \((-1)^k \Lambda_1^+, \gamma_1 L, (-1)^j \Lambda_2^+, \gamma_2, \gamma_3, (-1)^j L\) |
| \(D_5\) | \(k_2 > b_2\) | \(\Delta > 0\) | \(k_1 < b_1, k_1 + k_2 < 2b_1\) | 16 | \((-1)^k \Lambda_1^+, \gamma_1 L, (-1)^j \Lambda_2^+, \gamma_2, \gamma_3, (-1)^j L\) |
| \(C_1\) | \(\Delta = 0\) | \(k_1 + k_2 < 2b_1\) | \(k_1 + k_2 < 2b_1\) | 8 | \((-1)^k \Lambda_1^+, \gamma_1 L, (-1)^j \Lambda_2^+, \gamma_2, \gamma_3, (-1)^j L\) |
| \(C_2\) | \(\Delta = 0\) | \(k_1 + k_2 > 2b_1\) | \(k_1 + k_2 > 2b_1\) | 0 | \((-1)^k \Lambda_1^+, \gamma_1 L, (-1)^j \Lambda_2^+, \gamma_2, \gamma_3, (-1)^j L\) |
| \(D_4^1\) | \(\Delta < 0\) | \(\Delta < 0\) | \(\Delta < 0\) | 0 | \((-1)^k \Lambda_1^+, \gamma_1 L, (-1)^j \Lambda_2^+, \gamma_2, \gamma_3, (-1)^j L\) |

Table 2. Critical points for \(b_1 > b_2\). Here \(i, j, k \in \{0, 1\}\) and \(l = i + k - j\).

4.4. Classification of Critical points. Given the explicit computation of all the critical points on the smooth level surfaces, we can now classify all of them. Recall that the level surface \(S_c\) is defined by \((2.2)\). The tangent plane at \(p \in S_c\) is the kernel of the matrix \((3.1)\) formed by the gradients of the defining equations. Let \(p = (m_1, m_2, \gamma_1, \gamma_2, \gamma_3)\) be a critical point of the flow, then \(\gamma_3 = 0\) and none of the other coordinates vanishes. Thus, near a critical point \(p\) we have a local frame of the tangent space \(TS_c\) given by

\[
v_1 = \frac{b_1 \gamma_3}{m_1} \frac{\partial}{\partial m_1} - \frac{\gamma_3}{\gamma_1} \frac{\partial}{\partial \gamma_1} + \frac{\partial}{\partial \gamma_3} \quad \text{and} \quad v_2 = \frac{b_2 \gamma_2}{m_1} \frac{\partial}{\partial m_1} - \frac{b_2 \gamma_2}{m_2} \frac{\partial}{\partial m_2} - \frac{\gamma_2}{\gamma_1} \frac{\partial}{\partial \gamma_1} + \frac{\partial}{\partial \gamma_2}
\]
We see that integral curves of \( v_1 \) are given by \( \gamma_2 \equiv \text{const} \) and the integral curves of \( v_2 \) are given by \( \gamma_3 \equiv \text{const} \). In particular, \( \{ \gamma_2, \gamma_3 \} \) defines a local coordinate chart around \( p \). By an abuse of notation, we may write

\[
v_1 = \partial_{\gamma_3} \quad \text{and} \quad v_2 = \partial_{\gamma_2}
\]

Then the Suslov vector field on \( S_k \) near \( p \) can be written as

\[
X = (\gamma_2 m_2 - \gamma_1 m_1) \partial_{\gamma_3} - m_2 \gamma_3 \partial_{\gamma_2}
\]

Let \( P \) be a critical point of \( X \) and suppose that \( \gamma_2(P) = c \) and \( \gamma_3(P) = 0 \). From the equations (2.2), we compute that the linearization of \( X \) at \( P \) to be

\[
(4.4) \quad \left( m_2(P) - \frac{b_2 \gamma_2(P)^2}{m_2(P)} + \frac{m_1(P) \gamma_2(P)}{\gamma_1(P)} - \frac{b_1 \gamma_1(P) \gamma_2(P)}{m_1(P)} \right) (\gamma_2 - c) \partial_{\gamma_3} - m_2(P) \gamma_3 \partial_{\gamma_2}
\]

which gives the Jacobian of \( X \) at \( P \):

\[
J_X(P) = \begin{pmatrix}
0 & -m_2(P)
\frac{m_1(P) \gamma_2(P)}{\gamma_1(P)} - \frac{b_1 \gamma_1(P) \gamma_2(P)}{m_1(P)} & 0
\end{pmatrix}
\]

The characteristic polynomial of \( J_X(P) \) is

\[
\lambda^2 - m_2(P) \left( m_2(P) - \frac{b_2 \gamma_2(P)^2}{m_2(P)} + \frac{m_1(P) \gamma_2(P)}{\gamma_1(P)} - \frac{b_1 \gamma_1(P) \gamma_2(P)}{m_1(P)} \right)
\]

which simplifies to

\[
(4.5) \quad \lambda^2 + k_1 + k_2 - 2(m_1(P)^2 + m_2(P)^2)
\]

since at \( P \), we have \( \gamma_2(P)m_2(P) - \gamma_1(P)m_1(P) = 0 \). The type of the singularity is determined by the roots of the characteristic polynomial in (4.5).

First consider the case where \( b_1 = b_2 = b \). In this case, the flow has 8 critical points on the level set \( S_k \) when \( (k_1, k_2) \in D_2 \cup D_5 \), and no critical points in other regions.

**Proposition 4.3.** When \( b_1 = b_2 = b \), the critical points on \( S_k \) are all saddles if \( k \in D_2 \), and are all centers if \( k \in D_5 \).

**Proof.** In this case, the explicit coordinates for the singular points in Table 1 lead to

\[
m_1(P)^2 + m_2(P)^2 = k_1 + k_2 - b
\]

which implies that the roots of the characteristic polynomial is given by

\[
\pm \sqrt{2b - (k_1 + k_2)}
\]

The statement follows noticing that \( k_1 + k_2 < 2b \) in \( D_2 \), while \( k_1 + k_2 > 2b \) in \( D_5 \). \( \square \)

Next, consider \( b_1 \neq b_2 \). Without loss of generality, we suppose that \( b_1 > b_2 \).

**Proposition 4.4.** Suppose that \( b_1 > b_2 \). When \( S_k \) is a smooth 2-manifold, we have:

- If \( (k_1, k_2) \in D_2 \) then the 8 critical points are all saddles.
- If \( (k_1, k_2) \in D_4 \cup D_3^1 \) then there are 8 centers and 8 saddles.
- If \( (k_1, k_2) \in D_5 \) then the 8 critical points are all centers.
- If \( (k_1, k_2) \in C_1 \) then there are 8 non-hyperbolic critical points.
**Proof.** Using (2.2) and the fact that $\gamma_3(P) = 0$ at critical point $P$, we see that

$$m_1(P)^2 + m_2(P)^2 = k_1 + k_2 - b_2 - (b_1 - b_2)\gamma_1(P)^2$$

Thus (4.5) becomes

$$\lambda^2 - 2(b_1 - b_2) \left[ \frac{k_1 + k_2 - 2b_2}{2(b_1 - b_2)} - \gamma_1(P)^2 \right]$$

When $\Delta \neq 0$, the critical points are non-degenerate. Let’s call the critical points of the form $((-1)^i \Gamma_1^+ , (-1)^j \Gamma_2^+ , (-1)^k \Lambda_1^+ , (-1)^l \Lambda_2^+)$ the $+$-critical points, and the critical points of the form $((-1)^i \Gamma_1^- , (-1)^j \Gamma_2^- , (-1)^k \Lambda_1^- , (-1)^l \Lambda_2^-)$ the $-$-critical points. At the $\pm$-critical points, by (4.3), (4.5) further simplifies to

$$\lambda^2 \mp \sqrt{\Delta},$$

respectively.

In particular, all $+$-critical points are saddles and $-$-critical points are centers. This gives the first three statements.

When $(k_1, k_2) \in C_1$, we have $\Delta = 0$ at the critical points and they are all degenerate. The linearization (4.4) of $X$ at a critical point $P$ here becomes

$$-m_2(P)\gamma_3 \partial_{\gamma_3}$$

with $m_2(P) \neq 0$ which implies that they are nonhyperbolic. □

4.5. **Periodic orbits.** Recall that on each level surface $S_k$, an orbit of the Suslov flow projects to a portion of an orbit of a linear flow on the torus and the critical points of the Suslov flow correspond to precisely the points where $\partial U_k$ is tangent to the linear flow. Thus a generic orbit of the flow does not contain any critical point in its closure, and we say such generic orbits *non-critical.*

---

![Figure 6](image_url)

(0) $b_1 = 4$, $b_2 = 1$, $k_1 = 2$, and $k_2 = 3/4$ (b) $b_1 = 4$, $b_2 = 9$, $k_1 = 2$, and $k_2 = 27/4$

**Figure 6.** The flow for $k \in D_2$, for two different values of $k$.

Figures 6 and 7 illustrate the projection of the Suslov flow when $k \in D_2$ and $k \in D_4$, respectively. We notice that Figure 7a corresponds to $k \in D_4^4$, where there is no critical
point, while Figure 7b corresponds to $k \in D_4^1 \cup D_4^2$. One can understand the periodicity of the Suslov orbits from these projections.

**Lemma 4.5.** Let $O$ be a non-critical orbit of the Suslov flow on the level set $S_k$ and $O_T$ its projection to the torus. If $O_T \cap \partial U_k$ contains at least 2 points, then $O$ is periodic.

**Proof.** Let $p, q \in O_T \cap \partial U_k$ and let $\ell$ be the component of $O_T$ such that $\partial \ell = \{p, q\}$. By Lemma 3.3, we see that $(\varphi_k \circ p_k)^{-1}(\ell) \cong S^1$. Since $O$ is connected, we see that $O \cong S^1$, i.e. it is periodic. \qed

From the proof, we also see that the torus projection of the closure of a Suslov orbit can have at most two intersection points with $\partial U_k$. The following proposition states that in a large open set of the configuration space, Suslov orbits are generically periodic.

**Proposition 4.6.** Let $Q_1 = \bigcup_{k \notin \mathcal{T}} S_k$. Then Suslov orbits in $Q_1$ are generically periodic.

**Proof.** Suppose that $\sqrt{b_2/b_1} \in \mathbb{Q}$, then the corresponding linear flow on $T^2$ are periodic. Let $O$ be a Suslov orbit on an smooth level surface $S_k$, then it may not be periodic only if its torus projection $O_T$ contains a critical point in its closure. There is a finite number of those non-periodic orbits on each level surface, which implies that generic Suslov orbits are periodic. Note that in this case, we do not have to restrict to $Q_1$.

Suppose that $\sqrt{b_2/b_1} \notin \mathbb{Q}$, then the corresponding linear flow on $T^2$ are not periodic and we restrict the consideration to $k \in Q_1$. For such $k$, $\partial U_k \neq \emptyset$, and $T^2 \setminus U_k$ is an open subset. Any orbit of the corresponding linear flow is dense in $T^2$, and intersects $\partial U_k$ infinitely many times. Since there are only finitely many critical points on each level surface $S_k$, there are only finitely many linear orbits on $T^2$ that intersect with the torus projection of the critical points. By Lemma 4.5 all the Suslov orbits are periodic, except for a finite number which connects critical points. \qed
We remark that when there is no critical point on a level surface in $Q$, e.g. $k \in D_4^1$, all Suslov orbits on such $S_k$ are periodic. Furthermore, when a Suslov orbit is not periodic, it can be either homoclinic or heteroclinic, e.g. Figure 6a depicts 4 heteroclinic orbits and 8 homoclinic orbits, while in Figure 7b there are 16 homoclinic orbits.

4.6. **Topology of the level surfaces** $S_k$ via the Poincaré-Hopf theorem. The Poincaré-Hopf theorem [5] provides a deep link between a purely analytic concept, namely the index of a vector field, and a purely topological one, that is, the Euler characteristic. Recall that the Euler characteristic of a compact connected orientable two dimensional manifold is given by

$$\chi = 2 - 2g,$$

where $g$ is the genus, that is the number of “holes”, and that such manifold is determined, up to an homeomorphism, by its genus. The Poincaré-Hopf theorem allows us to determine the topology of $M$ by counting the indices of the zeroes of a vector field on $M$.

**Theorem 4.7** (Poincaré-Hopf). Let $M$ be a compact manifold and let $v$ be a smooth vector field on $M$ with isolated zeroes. If $M$ has a boundary, then $v$ is required to point outward at all boundary points. Then, the sum of the indices at the zeroes of such vector fields is equal to the Euler characteristic of $M$, that is, we have

$$\chi(M) = \sum \text{index}_{z_i}(v).$$

We now use the Poincaré-Hopf theorem to give an alternative proof of Proposition 3.5. Since on a compact two manifold the index of a sink, a source, or a center is $+1$, and the index of a hyperbolic saddle point is $-1$, the classification of the critical points given in Proposition 4.4 together with the knowledge of the number of connected components of the manifolds gives the proof for $\Delta \neq 0$. For instance, if $k \in D_2$, then there are 8 saddle points, so that $\chi(S_k) = -8$, and $g = \frac{2 - \chi}{2} = 5$. If $\Delta = 0$, the critical points are all degenerate and the vector field near the critical points is given by (4.6). In this case it is easy to see that the index of any critical point is 0, and thus $\chi(S_k) = 0$. Since there are two connected components $g = 1$ on each of them. It follows that $S_k$ is isomorphic to two copies of $T^2$.

In [2] a similar approach was used to obtain the topology of Suslov’s problem. The main difference is that the authors used an extension of the Poincaré-Hopf theorem that applies to compact manifolds with boundary even when the vector field does not point outward at all boundary points.

5. **Physical motion**

5.1. **Poisson sphere.** The 2-sphere $\gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1$ is known as the Poisson sphere. Let $\pi$ be the projection of $S_k$ onto the Poisson sphere. The *domain of possible motion* (DPM) corresponding to $k \in \mathbb{R}^2$ is the set $P_k = \pi(S_k) \subset S^2$, that is, it is the image of the projection of $S_k$ to the Poisson sphere [3]. If $p \in S^2$ is a point on the Poisson sphere, a vector $v \in \mathbb{R}^2$ such that $(p, v) \in S_k$ is said to be an *admissible velocity* at the point $p \in S^2$. A classification of the possible types of DPMs together with a study
of the set of admissible velocities gives a topological and geometrical description of
the mechanical system and it is useful in describing the main features of the physical
motion for various values of $k$. We rewrite the equations as

\begin{align*}
(1) \gamma_1^2 &= \frac{1}{b_1}(k_1 - m_1^2), \\
(2) \gamma_2^2 &= \frac{1}{b_2}(k_2 - m_2^2)
\end{align*}

and $\gamma_3^2 = 1 - \gamma_1^2 - \gamma_2^2$. The cardinality of the preimage of a point $x = (\gamma_1, \gamma_2, \gamma_3) \in S_k$ is
given by the number of pairs of $(m_1, m_2)$ that satisfy the equations (1) and (2) above.
Then $x \in P_k$ iff

\begin{align*}
\gamma_2 \leq \frac{k_1}{b_1} \quad \text{and} \quad \gamma_2 \leq \frac{k_2}{b_2}
\end{align*}

or equivalently we have

\begin{equation}
(5.1) \quad P_k = \{\gamma_1^2 + \gamma_2^2 + \gamma_3^2 = 1\} \cap \left\{(\gamma_1, \gamma_2, \gamma_3) \in \left[-\sqrt{\frac{k_1}{b_1}}, \sqrt{\frac{k_1}{b_1}}\right] \times \left[-\sqrt{\frac{k_2}{b_2}}, \sqrt{\frac{k_2}{b_2}}\right] \times \mathbb{R}\right\}
\end{equation}

In the interior of $P_k$, we have $m_1 \neq 0$ and $m_2 \neq 0$, which implies that the projection $S_k \to P_k$ is 4-to-1 in the interior.

The region $P_k$ may have boundary components, over which one or both of $m_1$ and $m_2$
vanish. If exactly one of $m_1$ and $m_2$ vanishes, the projection is 2-to-1. If $m_1 = 0$, then
$\dot{\gamma}_1 = 0$, and if $m_2 = 0$, then $\dot{\gamma}_2 = 0$. In the case $m_1 = m_2 = 0$, the corresponding points
in $P_k$ are corners and the projection is 1-to-1 and $\dot{\gamma}_1 = \dot{\gamma}_2 = \dot{\gamma}_3 = 0$. The diagrams
below illustrates the regions $P_k$ for various values of $k$.

Clear pictures emerge when the observations so far are combined. By (5.1), the
image $P_k$ of $S_k$ on the Poisson sphere is bounded by

\begin{align*}
\gamma_1 = \pm \sqrt{\frac{k_1}{b_1}} \quad \text{and} \quad \gamma_2 = \pm \sqrt{\frac{k_2}{b_2}}
\end{align*}

which correspond precisely to the following lines on the flat torus $T^2$, as indicated by
the dashed lines in Figure 3:

\begin{align*}
\theta_1 = \pm \frac{\pi}{2} \quad \text{and} \quad \theta_2 = 0 \text{ or } \pi
\end{align*}

The dashed lines divide $T^2$ into four components, and the projection $\pi : S_k \to P_k$
restricted to each component is 1-to-1; and the image of shaded region contained in each
of the components coincide. The following proposition provides a detailed classification
of the DPM for various values of $k \in \mathbb{R}^2$.

**Proposition 5.1.** Over the interior of the domain of possible motion $P_k$, the projection
$\pi : S_k \to P_k$ is 4-to-1. On the boundary components of $P_k$, the projection is 2-to-1,
except for over the corners when $k \in D_1$, where it is 1-to-1. Moreover, we have

1. For $k \in D_1$, each torus in $S_k$ is projected onto a component of $P_k$. Each
   component of $P_k$ is a square, see figure 8(1).
2. For $k \in D_2$ the set $P_k$ is a sphere with four holes as depicted in figure 8(2).
Figure 8. The domain of possible motion $P_k$ for various values of $k$. (1) For $k \in D_1$, $P_k$ has the form of two squares on opposite sides of the Poisson sphere. (2) For $k \in D_2$, $P_k$ is a sphere with four holes. (3) For $k \in D_3, D_4$, $P_k$ is a horizontal or vertical band wrapping around the sphere. (4) For $k \in D_5$, $P_k$ is the whole sphere.

(3) For $k \in D_3$ or $D_4$, the projection $\pi$ restricted to each torus component of $S_k$ is 2-to-1 in the interior of $P_k$. $P_k$ is a band wrapping around the Poisson sphere, see figure 8(3).

(4) For $k \in D_5$, the projection $\pi$ is an isomorphism when restricted to each $S^2$-component of $S_k$, see figure 8(4).

We can now use Proposition 5.1 to understand the physical motion of the rigid body. If $k \in D_1$, the trajectories in each component of $P_k$ are similar to Lissajous figures (the sum of independent horizontal and vertical oscillations). For each point inside each square there are four admissible velocities, there are two on its sides and one on the vertices. If $\sqrt{b_2/b_1} \not\in \mathbb{Q}$, then the trajectories is dense in the squares, otherwise they are periodic. In either case $E_3$ wobbles around the vertical direction, while $E_1$ remains close to horizontal and the wheels remain close to being vertical (see figure 1).

If $k \in D_2$, almost all the trajectories are periodic except for a finite number of orbits which connect critical points. For points in the interior of $P_k$ there are four admissible velocities. There are two admissible velocities on the boundary of $P_k$. This means that there are two trajectories for each point in the interior of $P_k$ and each trajectory can be followed in either direction. The physical motion in this case can be distinguished from the previous case since $E_3$ can go from pointing upward to pointing downward.

If $k \in D_3$, the trajectories are confined in a band wrapping around the sphere and alternatively touch the upper and the lower boundary of the band. Since this region is...
the image of two tori there are two admissible velocities at each point, and a point can move along the trajectories in either direction. In this case $E_3$ performs a complete revolution wobbling about the vertical plane spanned by $e_1$ and $e_3$. The wheels remain close to vertical (see Figure 1). The case $k \in D_4$ is similar. When $b_1 > b_2$, certain subregions of $D_4$ allow homoclinic or heteroclinic orbits. From Figure 7b, we see that in this case, the behaviour of periodic orbits changes drastically on either side of a homoclinic or heteroclinic orbit.

If $k \in D_5$, the trajectories are homeomorphic to circles. In this case there are four possible velocities for each point on $P_k$. It follows that there are two trajectories for each point on the Poisson sphere and each trajectory can be followed in either direction.

Acknowledgements

The authors wish to express their appreciation for helpful discussions with Luis Garcia-Naranjo and Dmitry Zenkov. The research was supported in part by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grants (SH, MS).

References

1. Yuri N Fedorov and Božidar Jovanović, Quasi-Chaplygin systems and nonholonomic rigid body dynamics, Letters in Mathematical Physics 76 (2006), no. 2, 215–230.
2. Oscar E Fernandez, Anthony M Bloch, and Dmitry V Zenkov, The geometry and integrability of the Suslov problem, Journal of Mathematical Physics 55 (2014), no. 11, 112704.
3. Anatolij T Fomenko, Visual geometry and topology, Springer Science & Business Media, 2012.
4. Valerii Vasil’evich Kozlov, On the integration theory of equations of nonholonomic mechanics, Regular and Chaotic Dynamics 7 (2002), no. 2, 161–176.
5. John Willard Milnor, Topology from the Differentiable Viewpoint, Princeton University Press, 1997 (en).
6. GK Suslov, Theoretical mechanics, Gostekhizdat, Moscow 3 (1946), 40–43.
7. Ya V Tatarinov, Construction of non-torical invariant manifolds in a certain integrable nonholonomic problem, Usp. Mat. Nauk 40 (1985), no. 5, 216 (russian).
8. ______, Separation of variables and new topological phenomena in holonomic and nonholonomic systems, Trudy Seminar po Vekt. i Tenz. Anal. 23 (1988), 160–174 (russian).
9. V Wagner, On the geometrical interpretation of the motion of nonholonomic dynamical systems, Trudy Seminar po Vekt. i Tenz. Anal. 23 (1941), no. 5, 301–327 (russian).