The Volume Entropy of Sol

Richard Evan Schwartz *

April 2, 2020

Abstract

Let Sol be the 3-dimensional solvable Lie group whose underlying space is $\boldsymbol{R}^3$ and whose left-invariant Riemannian metric is given by

$$e^{-2z}dx^2 + e^{2z}dy^2 + dz^2.$$

Building on previous joint work with Matei Coiculescu, which characterizes the cut locus in Sol, we prove that Sol has volume entropy 1, the same as that of the hyperbolic plane.

1 Introduction

Sol is one of the 8 Thurston geometries [**Th**], the one which uniformizes torus bundles which fiber over the circle with Anosov monodromy. Sol has been studied in various contexts: coarse geometry [**EFW**], [**B**]; minimal surfaces [**LM**] (etc.); its geodesics [**G**], [**T**], [**K**], [**BS**]; connections to Hamiltonian systems [**A**], [**BT**]; and finally virtual reality [**CMST**]. Our paper [**CS**] has a more extensive discussion of these many references.

In [**CS**], Matei Coiculescu and I give an exact characterization of which geodesic segments in Sol are length minimizers, thereby giving a precise description of the cut locus of the identity in Sol. As a consequence, we proved that the metric spheres in Sol are topological spheres, smooth away from 4 singular arcs. We will summarize the characterization in §3 and explain the main ideas in the proof.

*Supported by N.S.F. Grant DMS-1807320

Ian Agol recently pointed out to me that our exact characterization of the cut locus in Sol might help us determine the growth rate in balls in Sol, i.e., the *volume entropy*. This is the quantity

$$V(\text{Sol}) = \limsup_{r\to\infty} \frac{\log V_r}{r}. \tag{1}$$

Here V_n is the volume of the metric ball of radius n centered at the origin. Since Sol is a Lie group, equipped with a left invariant metric, we only have to consider the growth of balls which are centered at the origin to define the volume entropy. Also, the volume entropy does not change when we scale the metric by a fixed factor.

The hyperbolic plane $\boldsymbol{H}^2$ has volume entropy 1 because the metric ball of radius r has volume

$$H_r = 2\pi(\cosh r - 1) \sim \pi e^r \tag{2}$$

when the metric is scaled to have curvature -1. More generally, the volume entropy of $\boldsymbol{H}^n$ is $n-1$.

We equip Sol with the metric

$$e^{-2z}dx^2 + e^{2z}dy^2 + dz^2. \tag{3}$$

In this metric the planes $X = 0$ and $Y = 0$ have sectional curvature -1. Let V_r denote the volume of ball of radius n in Sol.

Theorem 1.1 (Volume Entropy) *There are constants C_1 and C_2 such that*

$$2e^r - 2 \leq V_r \leq C_1 e^r + C_2$$

In particular, Sol has volume entropy 1*.*

Our proof will show that $C_1 < 676$. Based on computer plots, I think that this is far from optimial. I don't have an estimate on C_2. This constant appears because we only analyze what happens for sufficiently large spheres, and we do not get an effective estimate on what "sufficiently large" means. However, we notice that all the relevant quantities seem to stabilize pretty quickly: one seems all the phenonena already when looking at a sphere of radius 8 in Sol. It seems plausible that our sharp characterization of the spheres in Sol could lead to an actual formula for the asymptotic volume, but I didn't try for this.

The idea of the proof is based on a simple observation. Let $\eta_X : \boldsymbol{R}^3 \to \boldsymbol{R}$ the projection into the plane $X = 0$. Likewise define η_Y and η_Z. If $T \subset \boldsymbol{R}^3$ is a planar triangle, then

$$\frac{\text{area}(t)}{\text{area}(\eta_W(T))} \leq \sqrt{3}$$

for some choice of $W \in \{X, Y, Z\}$. By symmetry, the same principle holds in Sol, provided that we replace $\sqrt{3}$ by $\sqrt{3}+\epsilon$ and we take T sufficiently small. We are still projecting onto the coordinate planes, but we are using the Sol metric to measure areas. Using polygonal approximation and continuity, we get the following result.

Lemma 1.2 *Let Σ be a smooth bounded surface in Sol. Then the area of Σ is at most*

$$\sqrt{3} \times (N_X A_X + N_Y A_Y + N_Z A_Z).$$

Here A_W is the area of $\eta_W(\Sigma)$ and N_W is the maximum number of pre-images of the plane $W = 0$ under η_W^{-1}. We proved that the spheres in Sol are smooth away from 4 analytic arcs, so Lemma 1.2 also applies to them. The rest of the paper is devoted to establishing the following *projection estimates:*

1. $A_X(\mathcal{S}_r) = A_Y(\mathcal{S}_r) = 2\pi(\cosh(r) - 1) < \pi e^r$.
2. $N_X(\mathcal{S}_r) = N_Y(\mathcal{S}_r) = 2$.
3. $A_Z(\mathcal{S}_r) < 96e^r$ for sufficiently large r.
4. $N_Z(\mathcal{S}_r) = 4$ for sufficiently large r.

Here $\mathcal{S}_r$ is the metric sphere of radius r in Sol. When we apply the result in Lemma 1.2 we see that $\mathcal{S}_r < 696e^r$ at least if r is sufficiently large. We get a similar bound on the volumes of balls by integrating this bound. The factor of 96 in Projection Estimate 3 is where we have some slop.

Projection Estimates 1 and 2 are straightforward given our description of the Sol spheres. Projection Estimate 3 is not too bad given the full force of the main result of [**CS**]. The hard work in the paper involves dealing with Projection Estimate 4. This requires an extensive investigation of the main ODEs associated to the geodesics in Sol, especially their asymptotic properties.

This paper is organized as follows.

- In §2 we prove some preliminary technical results that we use later in the proof. In particular, we quickly dispense with the lower bound in Theorem 1.1.

- In §3, we recall the main theorem from [**CS**] and sketch its proof. Following this, we deduce some additional useful structure that we will need for the projection estimates.

- In §4 we prove Projection Estimates 1 and 2.

- In §5 we prove Projection Estimates 3 and 4 modulo the detail that a certain curve in the plane is smooth and regular except at a single cusp. See Figure 5.3. Proving this result, which we call the Embedding Theorem, turns out to be a fight with a system of nonlinear ODEs. We introduce this ODE in §3.5.

- In §6 we prove the Embedding Theorem modulo a detail which we call the Monotonicity Lemma, a statement about the ODE from §3.5.

- in §7-8 we prove the Monotonicity Lemma. This is where all the ODE calculations come in.

I would like to thank Ian Agol for asking me about the volume entropy of Sol. I would also like to thank Justin Holmer, Anton Izosimov, Boris Khesin, Mark Levi, and Benoit Pausader, who all gave me advice about Lemma 7.7. (This is kind of an embarrassment of riches.) Finally, this paper owes a big intellectual debt to Matei Coiculescu, my co-author in [**CS**]. For one thing, Matei and I have been steadily talking about Sol for almost a year now and we developed a lot of the intuition for Sol together. For another thing, Equation 43, which Matei discovered and then proved, gives a very nice proof of one part of the Monotonicity Lemma.

2 Preliminaries

2.1 Basic Properties of Sol

The underlying space for Sol is $\boldsymbol{R}^3$. Again, the left invariant metric on Sol is given by

$$e^{-2z}dx^2 + e^{2z}dy^2 + dz^2. \tag{4}$$

The group law on Sol is

$$(x, y, z) * (a, b, c) = (e^z a + x, e^{-z} b + y, c + z). \tag{5}$$

Left multiplication is an isometry. We identify $\boldsymbol{R}^3$ with the Lie algebra of Sol in the obvious way. (See [**CS**, §2.1] if this does not seem obvious.)

Sol has 3 interesting foliations.

- The XY foliation is by (non-geodesically-embedded) Euclidean planes.
- The XZ foliation is by geodesically embedded hyperbolic planes.
- The YZ foliation is by geodesically embedded hyperbolic planes.

The complement of the union of the two planes $X = 0$ and $Y = 0$ is a union of 4 *sectors*. One of the sectors, the *positive sector*, consists of vectors of the form (x, y, z) with $x, y > 0$. The sectors are permuted by the Klein-4 group generated by isometric reflections in the planes $X = 0$ and $Y = 0$. The Riemannian exponential map E preserves the sectors. Usually, this symmetry will allow us to confine our attention to the positive sector.

Notation: For each $W \in \{X, Y, Z\}$, the plane Π_W is given by $W = 0$ and the map $\eta_W : \text{Sol} \to \Pi_W$ is the projection onto Π_X obtained by just dropping the W coordinate. We also let π_Z denote projection onto the Z-axis. Thus, $\pi_Z(x, y, z) = z$.

2.2 An Integral Formula

Let $\mathcal{S}_r$ and $\mathcal{B}_r$ respectively denote the metric sphere and ball in Sol of radius r centered at the origin. Our main result in [**CS**] implies that $\mathcal{S}_r$ is smooth away from 4 arcs of hyperbolas contained in the plane $Z = 0$. Hence, $\mathcal{S}_r$ has a well-defined surface area s_r. Recall that V_r is the volume of $\mathcal{B}_r$.

Lemma 2.1 $V_r = \int_0^r s_t dt$.

Proof: This holds in any Riemannian manifold, and the main point is that a minimizing geodesic segment joining the origin to a point $p \in \mathcal{S}_r$ is perpendicular to $\mathcal{S}_t$ for all $t < r$. ♠

Let us quickly dispense with the lower bound in Theorem 1.1. The projection η_X does not increase area. The half of $\mathcal{S}_r$ consisting of vectors of the form (x, y, z) with $x \geq 0$ projects onto the hyperbolic disk of radius r in the hyperbolic plane $X = 0$. Hence, $\mathcal{S}_r$ has area at least $4\pi(\cosh(r) - 1)$. By Lemma 2.1, we have

$$V_r \geq 4\pi(\sinh(r) - r) > 2e^r - 2.$$

2.3 The Disk Lemma

Here we recall a result from topology. This result will be useful, in §5, when we prove Projection Estimate 4.

Lemma 2.2 (Disk) *Let Δ be a topological disk. Let $h : \Delta \to \boldsymbol{R}^2$ be a map with the following properties:*

- *The restriction of h to the interior of Δ is a local diffeomorphism.*
- *The image $h(\partial\Delta)$ is a piecewise smooth curve having finitely many self-intersections.*

Given any $p \in \boldsymbol{R}^2 - h(\Delta)$, the number of preimages $h^{-1}(p)$ equals the unsigned number of times $h(\partial\Delta)$ winds around p.

Proof: This is a well-known result. Here we sketch the proof. Without loss of generality, we can assume that Δ is the unit disk in $\boldsymbol{R}^2$ and $h(0,0) \neq p$. Let Δ_s denote the disk of radius r centered at $(0,0)$. Also, we can assume that h is orientation preserving in the interior of Δ. Let $f(s)$ denote the number of times $h(\Delta_s)$ winds around p. For s near 0, we have $f(s) = 0$. The function f changes by ± 1 each time Δ_s crosses a point of $f^{-1}(p)$. The sign is always the same because h is orientation preserving. Hence the number of points in $f^{-1}(p)$ equals $f(1)$, up to sign. ♠

2.4 Elliptic Functions

Many of the quantities associated to Sol are expressed in terms of elliptic integrals. In this section we present some basic material on these special functions, and we also do two Mathematica calculations which we will need later on.

Basic Definitions: The complete elliptic functions of the first and second kind are given by

$$\mathcal{K}(m) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - m\sin^2\theta}}, \qquad \mathcal{E}(m) = \int_0^{\pi/2} \sqrt{1 - m\sin^2(\theta)}\; d\theta. \tag{6}$$

The first integral is defined for $m \in [0,1)$ and the second one is defined for $m \in [0,1]$. These functions satisfy the following differential equations. For a proof see any textbook on elliptic functions.

$$\frac{d\mathcal{K}}{dm} = \frac{(m-1)\mathcal{K} + \mathcal{E}}{2m - 2m^2}, \qquad \frac{d\mathcal{E}}{dm} = \frac{-\mathcal{K} + \mathcal{E}}{2m}. \tag{7}$$

Identities and Asymptotics: We will use the following classic identity.

$$\mathcal{K}(m) = \frac{\pi/2}{\mathrm{AGM}(\sqrt{1-m}, 1)}, \qquad m \in (0,1). \tag{8}$$

See [**BB**] for a proof. We also have

$$\mathcal{E}(1) = 1, \qquad \lim_{m \to 1} \frac{\mathcal{K}(m)}{|\log(1-m)|} = 2. \tag{9}$$

The second equation is much more than we need for the proof of Lemma 2.3 below. We just need to know that $\mathcal{K}(m)$ grows more slowly than $1/(1-m)$.

Two Calculations: Now we present the two semingly unmotivated calculations. They are related to Equation 17, which is fundamental to Sol. We will use these calculations in the proofs of Lemmas 7.3 and 8.1. The quantity Y_L below will be $Y_L(0)$ in Lemmas 7.3 and 8.1. We advise the reader to skip the proof of this result until reading Lemmas 7.3 and 8.1, because then the lemma will be motivated.

Define

$$L(y) = 4f(y)\mathcal{K} \circ g(y), \tag{10}$$

where

$$f(y) = \frac{1}{\sqrt{1 + 2y\sqrt{1-y^2}}} \quad g(y) = \frac{1 - 2y\sqrt{1-y^2}}{1 + 2y\sqrt{1-y^2}}. \tag{11}$$

Next define

$$Y_L = \frac{1}{y \times \frac{dL}{dy}}. \tag{12}$$

We use the notation $f_L \sim g_L$ if $\lim_{L\to\infty} f_L/g_L = 1$.

Lemma 2.3 *$Y_L \sim -1/2$ and $\frac{d}{dL}Y_L \sim 0$.*

Proof: We compute this in Mathematica, and we get a polynomial expression terms of $\sqrt{y}$ and $K = \mathcal{K} \circ g(y)$ and $E = \mathcal{E} \circ g(y)$. Taking the series expansions of the coefficients, we find that

$$Y_L = \frac{-1 - 5y + \cdots}{\Delta}, \quad \Delta = (2 + 12y + \cdots)E + (-4y - 24y^2 + \cdots)K. \tag{13}$$

Given that $g(y) \sim 1 - 4y$, we see from Equations 9 and 13 that $Y_L \sim -1/2$. Next we compute that

$$\frac{\partial}{\partial y}Y_L = \frac{(48y + \cdots)E + (-48y + \cdots)K}{\Delta^2}. \tag{14}$$

We see from Equations 9 and 14 that $\frac{\partial}{\partial y}Y_L \sim 0$. ♠

3 The Cut Locus Theorem: Synopsis

3.1 The Hamiltonian Flow

Let $G = \text{Sol}$. Let $S_1 \subset \boldsymbol{R}^3$ denote the unit sphere. at the origin in G. Given a unit speed geodesic γ, the tangent vector $\gamma'(t)$ is part of a left invariant vector field on G, and we let $\gamma^*(t) \in S_1$ be the restriction of this vector field to $(0,0,0)$. In terms of left multiplication on G, we have the formula

$$\gamma^*(t) = dL_{\gamma(t)^{-1}}(\gamma'(t)). \tag{15}$$

It turns out that γ^* satisfies the following differential equation.

$$\frac{d\gamma^*(t)}{dt} = \Sigma(\gamma^*(t)), \qquad \Sigma(x,y,z) = (+xz, -yz, -x^2+y^2). \tag{16}$$

This is explained one way in [**G**] and another way in [**CS**, §5.1]. (Our formula agrees with the one in [**G**] up to sign, and the difference of sign comes from the fact that our group law differs by a sign change from the one there.) This system in Equation 16 is really just geodesic flow on the unit tangent bundle of Sol, viewed in a left-invariant reference frame.

Let $F(x,y,z) = xy$. The flow lines of Σ lie in the level sets of F, and indeed Σ is the Hamiltonian flow generated by F. Most of the level sets of F are closed loops. We call these *loop level sets*. With the exception of the points in the planes $X = 0$ and $Y = 0$, and the points $(x,y,0)$ with $|x| = |y| = \sqrt{2}$, the remaining points lie in loop level sets.

Each loop level set Θ has an associated *period* $L = L_\Theta$, which is the time it takes a flowline – i.e., an integral curve – in Θ to flow exactly once around. Equation 17 below gives a formula. We can compare L to the length T of a geodesic segment γ associated to a flowline that starts at some point of Θ and flows for time T. We call γ *small*, *perfect*, or *large* according as $T < L$, or $T = L$, or $T > L$. In [**CS**, §5] we prove the following result:

Theorem 3.1 *Suppose $(x,y,z) \in S^2$ lies in a loop level set. Let $\alpha = \sqrt{|xy|}$. Then the period of the loop level set containing (x,y,z) is*

$$L_\alpha = \frac{\pi}{\text{AGM}(\alpha, \frac{1}{2}\sqrt{1+2\alpha^2})} = \frac{4}{\sqrt{1+2\alpha^2}} \times \mathcal{K}\left(\frac{1-2\alpha^2}{1+2\alpha^2}\right). \tag{17}$$

The second expression follows from the first expression and from Equation 8.

Each vector $V = (x,y,z)$ simultaneously corresponds to two objects:

- The flowline ϕ_V which starts at $V/\|V\|$ and goes for time $\|V\|$.
- The geodesic segment $\gamma_V = \{E(tV) |\ t \in [0,1]\}$.

Given a vector $V = (x, y, z)$ we define

$$\mu(V) = \text{AGM}(\sqrt{xy}, \frac{1}{2}\sqrt{(|x| + |y|)^2 + z^2}). \tag{18}$$

We call V *small*, *perfect*, or *large* according as $\mu(V)$ is less than, equal to, or greater than π. In view of Equation 17 here is what this means:

- If γ_V lies in the plane $X = 0$ or $Y = 0$ then V is small because $\mu(V) = 0$.
- If $V = (x, y, 0)$ where $|x| = |y|$ then V is small, perfect, or large according as $|x| < \pi$, $|x| = \pi$ or $|x| \geq \pi$.
- In all other cases, $V/\|V\|$ lies in a loop level set, and V is small, perfect, or large according as $\mu(V) < \pi$, $\mu(V) = \pi$, or $\mu(V) > \pi$.

3.2 The Main Result

Here is the main result from [**CS**].

Theorem 3.2 *Given any vector V, the geodesic segment γ_V is a distance minimizing geodesic if and only if $\mu(V) \leq \pi$. That is, γ_V is distance minimizing if and only if V is small or perfect. Moreover, if V and W are perfect vectors then $E(V) = E(W)$ if and only if $V = (x, y, z)$ and $W = (x, y, \pm z)$.*

Theorem 1.1 identifies the cut locus of the identity in Sol with the set of perfect vectors. The Riemannian exponential map E is a global diffeomorphism on the set of small vectors. Also, E is generically 2-to-1 on the set of perfect vectors. We will explain this last fact below.

Theorem 1.1 leads to a good description of the Sol metric sphere $\mathcal{S}_r$ of radius r. Let S_r denote the Euclidean sphere of radius r centered at the origin of $\boldsymbol{R}^3$. Let

$$S_r' = \mu^{-1}[0, \pi] \cap S_r. \tag{19}$$

The space S_r' is a 4-holed sphere. The boundary $\partial S_r'$, a union of 4 loops, is precisely the set of perfect vectors contained in S_r. Each of these loops is homothetic to one of the loop level sets on the unit sphere S^2. The Klein-4 symmetry explains why there are 4 such loops.

It follows from the Main Theorem that $\mathcal{S}_r = E(S'_r)$ and that E is a diffeomorphism when restricted to $S'_r - \partial S'_r$. On $\partial S'_r$, the map E is a 2-to-1 folding map which identifies partner points within each component. Thus, we see that $\mathcal{S}_r$ is obtained from a 4-holed sphere by gluing together each boundary component (to itself) in a 2-to-1 fashion. This reveals $\mathcal{S}_r$ to be a topological sphere which is smooth away from the set $E(\partial S'_r)$.

Let us say a bit more about the singular set $E(\partial S'_r)$. Let Π_Z denote the plane $Z = 0$. By Item 2 of §3.3, we have $E(V) \subset \Pi_Z$ when V is perfect. In particular, the singular set $E(\partial S_r)$ lies in Π_Z. More precisely, Items 2 and 4 from §3.3 below imply that the singular set consists of 4 arcs of hyperbolas contained in Π_Z.

3.3 Concatenation

In the next several sections, we outline the proof of Theorem 3.2. Our exposition here is an abbreviated version of what appears in [**CS**].

Given a (finite) flowline g we write $g = a|b$ if g is the concatenation of flowlines a and b. That is, a is the initial part of g and b is the final part. We call g *symmetric* if the endpoints of g have the form (x, y, z) and $(x, y, -z)$.

Let Λ_g denote the endpoint of the geodesic segment associated to g, when this geodesic segment starts at the origin. It follows from left-invariance of the metric that

$$\Lambda_g = \Lambda_a * \Lambda_b. \tag{20}$$

Since the third coordinates of elements of Sol commute, we have

$$\pi_Z(\Lambda_g) = \pi_Z(\Lambda_a) + \pi_Z(\Lambda_b). \tag{21}$$

Here $\pi_Z(x, y, z) = z$. More formally, π_Z is the quotient map from Sol to the quotient Sol/Π_Z. Here Π_Z is not just a Euclidean plane in Sol but also a maximal normal subroup. The integral form of Equation 21 is

$$\pi_Z(\Lambda_g) = \int_0^T z(t)\ dt. \tag{22}$$

Here we have set $g = (x, y, z)$, and T is the total time that g takes to get from start to finish.

These equations have a variety of consequences, which we work out in detail in [**CS**, §2].

1. If g is a small flowline then g is symmetric if and only if $\pi_Z(\Lambda_g) = 0$. Moreover, the geodesic segment corresponding to a small symmetric flowline only intersects Π_Z at its endpoints.

2. If g is a perfect flowline then $\pi_Z(\Lambda_g) = 0$. This follows from the fact that $g = a|b$ where a and b are both small symmetric.

3. If $V_\pm = (x, y, \pm z)$, then V_+ is perfect if and only if V_- is perfect. Furthermore $E(V_+) = E(V_-)$. This is because the corresponding flowlines g_+ and g_- can be written as $g_+ = a|b$ and $g_- = b|a$ where a and b are both small symmetric. But then Λ_a and Λ_b are horizontal translations in Sol and hence commute. Hence $\Lambda_{g_+} = \Lambda_{g_-}$. We call V_+ and V_- *partners*.

4. Suppose V_1 and V_2 are perfect vectors such that $V_1/\|V_1\|$ and $V_2/\|V_2\|$ lie in the same loop level set. Let $E(V_i) = (a_i, b_i, 0)$. We call $\sqrt{a_i b_i}$ the *holonomy* of V_i. Letting g_1 and g_2 be the corresponding flowlines, we can write $g_1 = a|b$ and $g_2 = b|a$ where a and b are both small. But then $\Lambda_{g_1} = (a_1, b_1, 0)$ and $\Lambda_{g_2} = (a_2, b_2, 0)$ are conjugate in Sol. This gives $a_1 b_1 = a_2 b_2$. Hence V_1 and V_2 have the same holonomy.

5. Given $V = (x, y, z)$ we define $\sigma(V) = y/x$. We prove that if V is a perfect vector, then $\sigma(E(V)) = 1/\sigma(V)$. We call this the Reciprocity Lemma. The proof is a more subtle working out of the consequences of the conjugacy idea discussed in Item 4.

3.4 Outline of the Proof

With these preliminaries out of the way, we turn directly to the proof of Theorem 3.2. Item 3 in §3.3 shows that the perfect geodesic segments corresponding to vectors of the form (x, y, z) where $z \neq 0$ are not unique distance minimizers. It also follows from Item 3 that perfect geodesics segments corresponding to vectors of the form $(x, y, 0)$ have conjugate points. Hence, large geodesic segments cannot be distance minimizers. This essentially proves half of Theorem 3.2.

The second half of Theorem 3.2, the converse, says that a small or perfect geodesic segment is a distance minimizer. Since every small geodesic segment is contained in a perfect geodesic segment, it suffices to prove that perfect geodesic segments are distance minimizers.

We first prove [**CS**, Corollary 2.10]: The map E is injective on the set of perfect vectors with positive coordinates. This step has 2 ideas. We first show (following [**G**]) that the holonomy is a monotone function of the loop level set. So, if $E(V_1) = E(V_2)$ then $V_1/\|V_1\|$ and $V_2/\|V_2\|$ lie in the same loop level set. We also have $\sigma(V_1) = \sigma(V_2)$, by Item 5 above. This forces $V_1 = V_2$.

We finish the proof by showing that if V is perfect and W is small then it is impossible for $E(V) = E(W)$. This is really the heart of [**CS**]. The argument involves the system of nonlinear ODEs we introduce in §3.5. It will turn out that the argument in this paper involves a deeper study of these same ODEs.

Let us go back to the argument. By symmetry, we can restrict ourselves to the case when V and W both lie in the positive sector. Let M and ∂M respectively denote the set of small and perfect vectors. We show that $E(\partial M)$ is contained in a subset $\partial N \subset \Pi_Z$. The boundary of ∂_N, which we denote by $\partial_0 N$, is the graph of a smooth function in polar coordinates. The yellow region in Figure 3.1 shows part of the portion of ∂N that lies in the positive sector. See Figure 5.1 for an expanded view. There 3 symmetrically placed components in the other sectors which we are not showing.

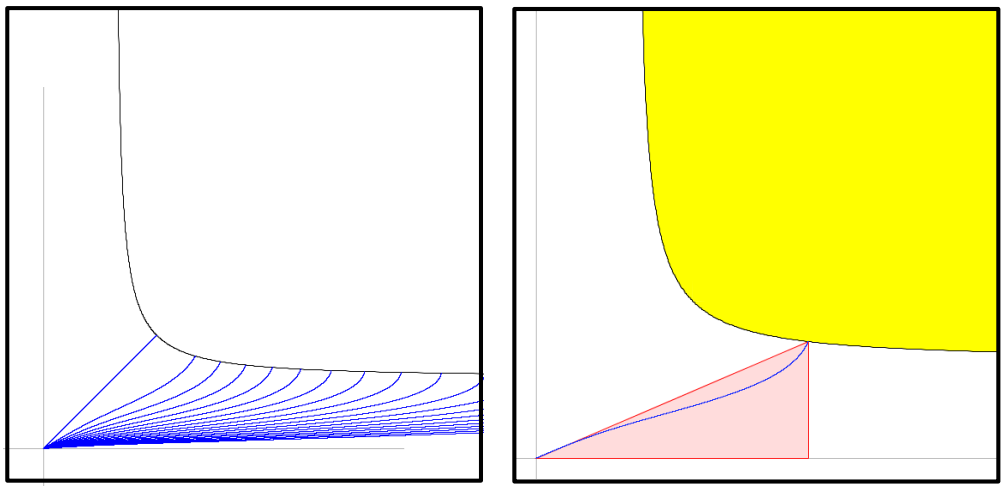

Figure 3.1: $\partial_0 N_+$ (black), ∂N_+ (yellow), Λ_L (blue), and Δ_L (red).

If we suppose that W is small and $E(W) = E(V)$ then the flowline corresponding to W must be small symmetric. We can arrange all the small symmetric flowlines in a given loop level set into two curves. One of the curves corresponds to small symmetric flowlines whose initial endpoint has

positive Z-coordinate. Given the loop level set of period L in the positive sector, we let Λ_L denote the image, under E, of the corresponding vectors. The blue curves in Figure 3.1 show Λ_L for various choices of L.

On the right side of Figure 3.1 we focus on Λ_5. We also draw the right triangle Δ_5 whose endpoints are the endpoints of Λ_5. We define the triangle Δ_L for other values of L in the same way. In [**CS**, §3] we prove that $\Lambda_L \subset \Delta_L$ and that the interior of Λ_L lies in the interior of Δ_L. Finally, we show that $\partial_0 N$ intersects Δ_L only at the top vertex. These ingredients combine to show that $\Lambda_L \cap \partial N = \emptyset$, and this shows that $E(V)$ and $E(W)$ cannot be equal.

3.5 The Differential Equation

We will now go into more detail about how the Bounding Triangle Theorem is proved. Let $\ell = L/2$. We consider the *backwards flow* along the structure field Σ, namely

$$x' = -xz, \qquad y' = +yz, \qquad z' = x^2 - y^2, \tag{23}$$

with initial conditions $x(0) > y(0) > 0$ and $z(0) = 0$ chosen so that the point is in the loop level set of period L. (We will often denote these functions as x_L, etc.) We let g_t be the small symmetric flowline whose endpoints are $(x(t), y(t), z(t))$ and $(x(t), y(t), -z(t))$. Then

$$\Lambda_L(t) = (a(t), b(t), 0) = \Lambda_{g_t}.$$

Taking the derivative, we have

$$(a', b', 0) = \Lambda_L'(t) = \lim_{\epsilon \to 0} \frac{\Lambda_L(t+\epsilon) - \Lambda(t)}{\epsilon},$$

$$\Lambda_L(t+\epsilon) \approx (\epsilon x, \epsilon y, \epsilon z) * (a, b, 0) * (\epsilon x, \epsilon y, -\epsilon z).$$

The approximation is true up to order ϵ^2 and $(*)$ denotes multiplication in Sol. A direct calculation gives

$$a' = 2x + az, \qquad b' = 2y - bz. \tag{24}$$

The initial conditions are $a(0) = b(0) = 0$. (We will often denote these functions as a_L and b_L.

Lemma 3.3

$$\frac{b(0)}{a(0)} = \frac{b(\ell)}{a(\ell)}. \tag{25}$$

Proof: This comes from L'hopital's rule and the Reciprocity Lemma. Let us take the opportunity to give a swift proof here. (This is another proof of the Reciprocity Lemma in a special case.) We have $(ax)' = 2x^2$ and $(by)' = 2y^2$. Also $a(0) = b(0) = 0$. Hence

$$a(\ell)x(\ell) = \int_0^\ell 2x^2\ dt, \qquad b(\ell)y(\ell) = \int_0^\ell 2y^2\ dt.$$

But these two integrals are equal, by symmetry. Hence $a(\ell)x(\ell) = b(\ell)y(\ell)$. Finally, we have $x(0) = y(\ell)$ and $y(0) = x(\ell)$ by symmetry. Combining these equations gives the result. ♠

The function $b(t)/a(t)$ has the same value at $t = 0$ and $t = \ell$. To finish the proof, we just have to show that $b(t)/a(t)$ cannot have a local maximum. This boils down to the fact that $ab'' - ba'' = 2ab(y^2 - x^2)$, a quantity which is negative for $t < \ell/2$ and positive for $t > \ell/2$. These properties, together with the fact that $a' > 0$, force $\Lambda_L \subset \Delta_L$. See [**CS**, §3] for more details.

We now derive an identity that does not appear in [**CS**]. Let $\underline{g}_t$ denote the first half of the flowline g_t; it connects the initial point of g_t to the midpoint of g_t. Say that the coordinates of $\Lambda_{\underline{g}_t}$ are $(\underline{a}(t), \underline{b}(t), \underline{c}(t))$. The coordinate $\underline{c}(t)$ is typically nonzero, but we do not care about it. Define

$$\underline{\Lambda}_L(t) = (\underline{a}_L(t), \underline{b}_L(t)) \subset \boldsymbol{R}^2. \tag{26}$$

In the next lemma we identity Π_Z with $\boldsymbol{R}^2$.

Lemma 3.4 (Doubling) $\underline{\Lambda}(t) = \frac{1}{2}\Lambda_L(t)$.

Proof: We have

$$(\underline{a}', \underline{b}', \underline{c}') = \lim_{\epsilon \to 0} \frac{\underline{\Lambda}_L(t+\epsilon) - \underline{\Lambda}(t)}{\epsilon}, \qquad \underline{\Lambda}_L(t+\epsilon) \approx (\epsilon x, \epsilon y, \epsilon z) * (a, b, c).$$

Taking the limit, we find that

$$\underline{a}' = z + \underline{a}x, \qquad \underline{b}' = z - \underline{b}x. \tag{27}$$

(Also $\underline{c}' = z$. We have the same initial conditions $\underline{a}(0) = \underline{b}(0)$ as above. Now notice that this solution to this equation is given by $\underline{a} = a/2$ and $\underline{b} = b/2$. ♠

4 The Hyperbolic Projections

4.1 A Picture

Recall that $\eta_X : \mathrm{Sol} \to \Pi_X$ is the orthogonal projection onto the plane $X = 0$. Figure 4.1 shows the projection of (part of) the positive sector of the sphere $\mathcal{S}_5$ into the plane Π_X. The smooth part of this sphere has a foliation by the images of the loop level sets under the Riemannian exponential map E. The grey curves are the projections of this foliation into Π_X. The black line segment is the projection of the set of singular points.

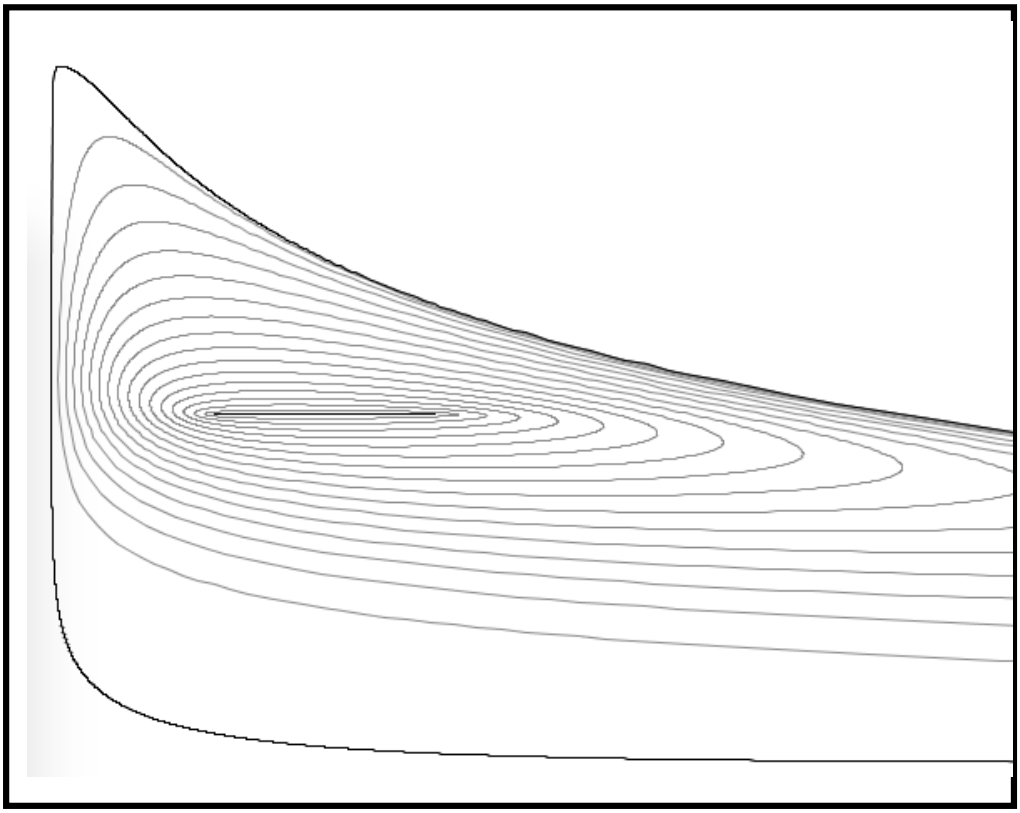

Figure 4.1: Projection into the plane Π_X.

It appears from the picture that the restriction of η_X to this sector is a homeomorphism onto its image. We will prove this result below. For convenience we take $r > \pi\sqrt{2}$.

4.2 Area Bound

In this section we prove Projection Estimate 1, listed at the end of §1.2. That is, we prove that $A_X(\mathcal{S}_r) < \pi e^r$ and $A_Y(\mathcal{S}_r) < \pi e^r$. By symmetry, it suffices to prove the result for the projection η_X into the plane Π_X.

Let $\mathcal{S}_r^+$ denote the subset of $\mathcal{S}_r$ consisting of points (x, y, z) with $x \geq 0$. Let $\mathcal{H}_r$ denote the hyperbolic disk of radius r contained in the plane Π_X and centered at the origin. All the points in the interior of $\mathcal{S}_r^+$ lie in the open positive sector. Because Π_X is a totally geodesic plane in Sol, we have $\partial\mathcal{S}_r^+ = \partial\mathcal{H}_r$.

Lemma 4.1 *η_X maps the interior of $\mathcal{S}_r^+$ into the interior of $\mathcal{H}_r$.*

Proof: If this is false, then there is a geodesic segment γ of length r, connecting $(0, 0, 0)$ to some point $p \in \mathcal{S}_r^+$ which remains entirely in the positive sector except for its initial point, $(0, 0, 0)$. The projection map η_X is distance non-increasing, and locally distance decreasing on any curve whose tangent vector is not in a plane of the form $X = \text{const.}$. This means that $\eta_X(\gamma)$ is shorter than γ. But then $\eta_X(\gamma)$ cannot reach the point $\eta_X(p) \in \partial\mathcal{H}_r$. ♠

Lemma 4.1 immediately gives us our estimate on $A_X(\mathcal{S}_r)$.

4.3 Multiplicity Bound

In this section we prove Projection Estimate 2, listed at the end of §1.2. That is, we prove that $N_X(\mathcal{S}_r) = 2$ and $N_Y(\mathcal{S}_r) = 2$. As above, it suffices to prove this result for the projection η_X into the plane Π_X.

To prove that $N_X(\mathcal{S}_r) = 2$ is suffices, by symmetry to show that η_X is an injective map from $\mathcal{S}_r^+$ to Π_X. The basic strategy is to show that η_X is locally injective. We also know that η_X is the identity on the boundary of $\mathcal{S}_r^+$, which already lies in the plane $X = 0$. (It is the boundary of the hyperbolic disk on Π_X of radius r centered at the origin.) Our injectivity result then follows from the Disk Lemma in §2.

For convenience we take $r > \pi\sqrt{2}$ in the next result, so that we don't have to discuss several cases. (The sphere $\mathcal{S}_r$ is smooth for $r < \pi\sqrt{2}$ and has 4 singular arcs for $r > \pi\sqrt{2}$.) The set of smooth points of $\mathcal{S}_r$ is a union of 4 open "punctured" disks. In each case, we are removing an analytic arc from an open topological disk and what remains is smooth. Figure 4.1 shows (a portion of) the η_X-projections of the smooth points of $\mathcal{S}_r$.

Lemma 4.2 *The differential $d\eta_X$ is injective at all the smooth points in the interior of $\mathcal{S}_r^+$*

Proof: Let S_r denote the subset of the sphere of radius r centered at the origin in the Lie algebra. As in the previous chapter, let S_r' denote the subset of S_r consisting of vectors which are either small or perfect. Let $p \in S_r'$ be some point. We think of p as a vector, so that $E(p) \in \mathcal{S}_r^+$. Let T_p be the tangent plane to S_r' at p. Let N_p be the unit normal to T_p. Since the perfect geodesic segments are minimizers, the small geodesic segments are unique minimizers without conjugate points. So, at the corresponding points of S_r', the differential dE_p is an isomorphism. We just have to show that $(1,0,0) \notin dE_p(T_p)$. We will suppose that $(1,0,0) \in dE_p(T_p)$ and derive a contradiction.

If $(1,0,0) \in dE_p(T_p)$, then the first component of $dE_p(N_p)$ is 0, because $dE_p(N_p)$ and $dE_p(T_p)$ are perpendicular. Let γ_p be the geodesic segment corresponding to p. The vector $dE_p(N_p)$ is the unit vector tangent to γ_p at its far endpoint – i.e., the endpoint not at the origin. This vector lies in the same left invariant vector field as the endpoint U_p of the flowline corresponding to p. If the first coordinate of U_p is 0, then the entire flowline lies in the plane $X = 0$. But then $E(p) \in \partial\mathcal{S}_r^+$. This is a contradiction. ♠

Lemma 4.3 *The map η_X is locally injective at each singular point of $\mathcal{S}_r^+$.*

Proof: As we showed in [**CS**], the singular set in $\mathcal{S}_r$ consists of 4 arcs of hyperbolas, each contained in the plane Π_Z. Each of these arcs lies in the interior of a different sector and is an arc of a hyperbola. These hyperbolas are all graphs of functions. The restriction of η_X to each hyperbola is therefore injective. We still need to see, however, that η_X is injective in neighborhoods of these singular sets, and not just on the singular sets. There are two cases.

Case 1: Consider a point p in the interior of the singular set in $\mathcal{S}_r^+$. By symmetry it suffices to consider the case when p is in the positive sector. The point p lies in the plane Π_Z and has its first two coordinates positive. There are exactly 2 points $p_+, p_- \in S_r'$ such that $E(p_+) = E(p_-) = p$. These points have the form $p_1 = (x,y,z)$ and $p_- = (x,y,-z)$. We called such points *partners*. In [**CS**, Lemma 2.8] we showed that $dE_{p_\pm}$ is non-singular. This crucially uses the fact that $p_\pm$ is a perfect vector whose third coordinate is

nonzero. The same argument as in the previous lemma now shows that the linear map $\eta_X \circ dE_{p_\pm}$ is an isomorphism from the tangent plane T_{p_+} to $\boldsymbol{R}^2$. But then $\eta_X \circ E$ is a diffeomorphism when restricted to an open neighborhood $U_\pm$ of $p_\pm$ in S'_r. The sets $U_\pm$ are disks with some of their boundary included. The portion of the included boundary consists of the perfect vectors in $\partial S'_r$ near $p_\pm$. See Figure 4.2.

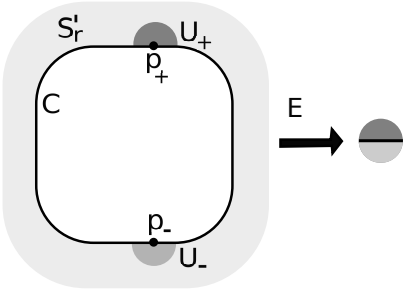

Figure 4.2: The neighborhoods U_+ and U_-.

Let C be the component of $\partial S'_r$ which contains p_+ and p_-. The image $E(U_+ - C)$ lies entirely below the plane Π_Z because the flowlines corresponding to vectors in $U_+ - C$ nearly wind the entirely around their loop level set but omit a small arc near p_+. Likewise, the image $E(U_- - C)$ lies entirely above the plane Π_Z. Hence $\eta_X \circ E(U_+ - C)$ and $\eta_x \circ E(U_- - C)$ are disjoint. Combining this what we know, we see that η_X is a homeomorphism in a neighborhood of $p \in \mathcal{S}_r$.

Case 2: Suppose that p is one of the endpoints of the singular set. This case is rather tricky to check directly. Suppose that there is some other point $q \in \mathcal{S}_n$ such that $\eta_X(p) = \eta_X(q)$. By Lemma 4.1, the point $\eta_X(p)$ is disjoint from the the hyperbolic circle $\mathcal{S}_r^+ \cap \Pi_X$. Hence q lies in the interior of $\mathcal{S}_r^+$. Since η_X is injective on the singular set, q must be a smooth point.

Since $\eta_X(q) = \eta_X(p)$ and $p \in \Pi_Z$, we have $q \in \Pi_Z$. This means that q corresponds to some small symmetric flowline. The point q is contained in a maximal connected arc $\mathcal{A} \subset \mathcal{S}_r^+$ consisting entirely of points corresponding to small symmetric flowlines. One endpoint of $\mathcal{A}$ is p. The other endpoint lies in the plane $X = 0$. The point q lies somewhere in the interior of $\mathcal{A}$. The map η_X sends $\mathcal{A}$ into the line $X = Z = 0$ and from Lemma 4.3, the restriction of η_X to $\mathcal{A}$ is locally injective. But a locally injective map from an arc into a line is injective. This contradicts the fact that $\eta_X(p) = \eta_X(q)$. ♠

Let D denote the quotient $S_r^+/\sim$ where the equivalence relation $\sim$ glues together partner points on the set of perfect vectors in S_r^+. The space D is a topological disk, and $h = \eta_X \circ E$ gives a map from D to Π_X. Combining Lemmas 4.2 and 4.3 we see that h is locally injective at each interior point of D. Moreover, $h(\partial D)$ is an embedded loop, just the boundary of a hyperbolic disk in Π_X. By the Disk Lemma, $h : D \to \eta_X$ is injective. But E is a bijection from D to $\mathcal{S}_r^+$. Hence $\eta_X : \mathcal{S}_r^+ \to \Pi_X$ is injective, as desired.

This completes the proof that $N_X(\mathcal{S}_r) = 2$.

5 The Euclidean Projection

5.1 The Cut Locus Image

Figure 5.1 shows a more of the yellow region in Figure 3.1.

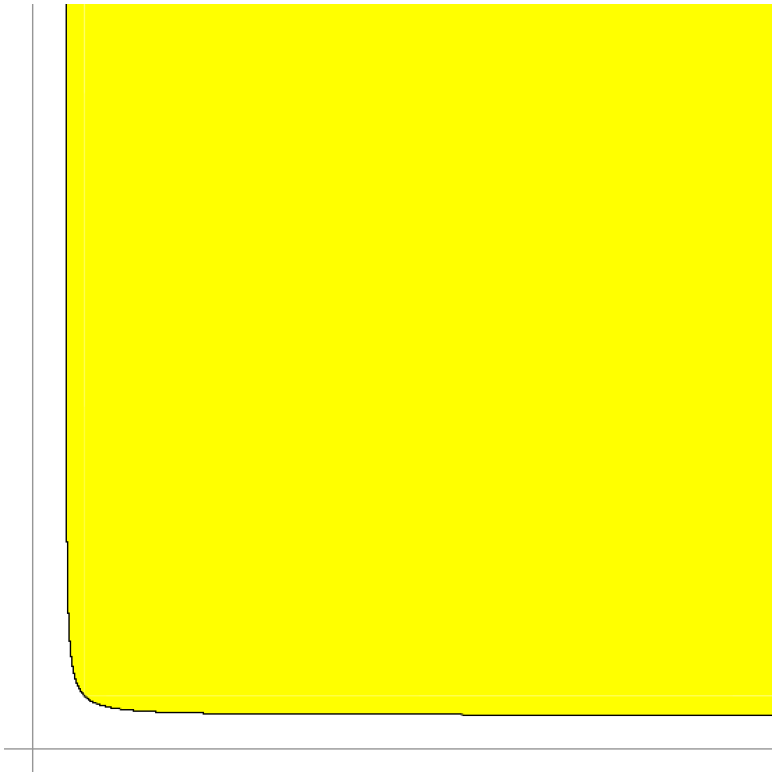

Figure 5.1: ∂N in the positive quadrant.

We can see from the picture that the boundary of the yellow curve has horizontal and vertical lines as asymptotes. We prove in this section that these asymptotes are the lines $X = 2$ and $Y = 2$. Given functions f and g of L we write $f \sim g$ is $f/g \to 1$ as $L \to \infty$. Our claims about the asymptotes follows from the statement that $\underline{a}_L(L) \sim 2$. Here $\underline{a}_L$ is as in Equation 26.

Lemma 5.1 $\underline{a}_L(L) \sim 2$.

Proof: We have $(\underline{a}x)' = x^2$ and $(\underline{b}y') = y^2$. Also $\underline{a}(0) = \underline{b}(0) = 0$. Therefore

$$\underline{a}(L)x(L) = \int_0^L x^2\ dt, \qquad \underline{b}(L)y(L) = \int_0^L y^2\ dt. \tag{28}$$

The two integrals are equal, by symmetry, so that $\underline{a}(L)x(L) = \underline{b}(L)y(L)$. Since $x(L) = x(0) \sim 1$, it suffices to prove that

$$\int_0^L x^2 \, dt \sim 2.$$

The function $t \to x^2 + y^2$ is periodic, with period $L/4$. Hence

$$\int_0^L x^2 \, dt = 2\int_0^{L/4} (x^2 + y^2)dt = 2\int_0^{L/4} (x^2 - y^2)dt + 2\int_0^{L/4} y^2 dt. \tag{29}$$

Now observe that

$$\int_0^{L/4} (x^2 - y^2)dt = \int_0^{L/4} z' \, dt = z(L/4) - z(0) = z(L/4) \sim 1. \tag{30}$$

To finish the proof, it suffices to show that

$$\int_0^{L/4} y^2 dt \sim 0.$$

Let α be such that $(x(0), y(0), 0)$ lies in the same loop level set as

$$(\alpha, \alpha, \sqrt{1 - 2\alpha^2}).$$

Then $y \leq \alpha$ on $[0, L/4]$ because y is monotone increasing on this interval. Hence, by Equation 17 and some algebraic manipulation,

$$\int_0^{L/4} y^2 dt \leq L_\alpha \times \alpha^2 = 2\alpha\sqrt{\frac{4\alpha^2}{1 + 2\alpha^2}} \times \mathcal{K}\left(\frac{1 - 2\alpha^2}{1 + 2\alpha^2}\right).$$

Setting $m = \frac{1-2\alpha^2}{1+2\alpha^2}$, we see that

$$\int_0^{L/4} y^2 dt \leq 2\alpha\sqrt{1 - m} \times \mathcal{K}(m). \tag{31}$$

As $L \to \infty$ we have $\alpha \to 0$ and $m \to 1$ and $\mathcal{K}(m) \sim -\log(1 - m)/2$. Hence, the right hand side of Equation 31 tends to 0 as $L \to \infty$. ♠

Say that a vector is *positive* if all its coordinates are positive. Say that a vector is *tiny* if its corresponding flowline is contained in a small symmetric flowline. In particular, tiny vectors are small.

Theorem 5.2 (Asymptotic) *Let $\epsilon > 0$ be given. Then the following is true for all sufficiently large L. If $V = (x, y, z)$ is a tiny positive vector whose corresponding flowline is contained in a loop level set of period L, then $E(V) = (a, b, 0)$ has the property that $b \in (0, 2 + \epsilon)$.*

Proof: This is really a consequence of Lemma 5.1. We first prove the result for a positive vector V corresponding to a small symmetric flowline. By the Bounding Triangle Theorem, $b_L(t) < b_L(L/2)$ for all $t \in (0, L/2)$. But $b_L(L/2) = \underline{a}_L(L)$, and we have $\underline{a}_L(L) \sim 2$ by Lemma 5.1. Hence for any $\epsilon > 0$ we can choose L large enough so that $b_L(t) \in (0, 2+\epsilon)$. This completes the proof in the special case we are considering.

Now consider the case when V is an arbitrary tiny positive vector. There is some $\lambda \geq 1$ so that λV corresponds to a small symmetric flowline. The X and Y coordinates of the curve $t \to E(tV)$ are increasing functions because it is impossible for the geodesic associated to V to be tangent to the hyperbolic foliations of Sol. (Otherwise this geodesic would be trapped inside a leaf of the foliation for all time.) In particular, the second coordinate of $E(V)$ is less or equal to the second coordinate of $E(\lambda V)$, which is in turn less than $2 + \epsilon$ by the case just considered. ♠

5.2 The Area Bound

Figure 5.2 shows the projection of (part of the positive sector of) $\mathcal{S}_5$ into the plane Π_Z. The small black arc of a hyperbola is the projection of the singular set. The outer black curve is

$$\min(xy^2, x^2y) = e^5.$$

Setting $(a, b, c) = E(V)$ for a small or perfect vector V of length 5, the picture suggests that $\min(a^2b, ab^2) < e^5$.

Lemma 5.3 below, the main ingredient in the Projection Estimate 3 listed at the end of §1.2, gives a weaker estimate along these lines which works at least for large radii spheres. Lemma 5.3 (which is just for the positive sector) will immediately imply that $A_Z(\mathcal{S}_r) < 96e^r$ provided that r is sufficiently large.

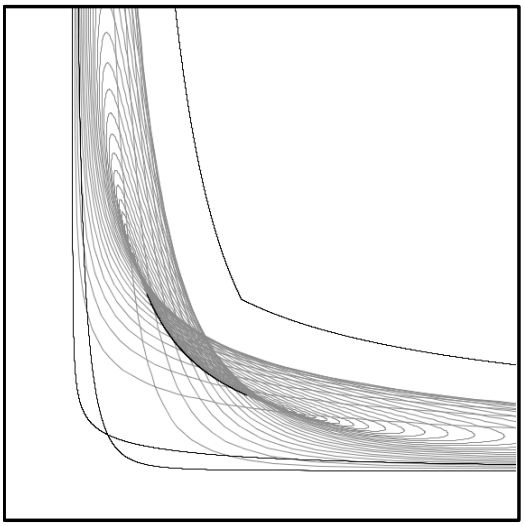

Figure 5.2: Projection into the plane η_Z.

Lemma 5.3 *Let V be a small or perfect positive vector of length r. Let L be the period of the loop level set containing $V/\|V\|$. Let*

$$(a, b, c) = E(V).$$

If L is sufficiently large then $\min(ab^2, a^2b) < 28e^r$ *and* $\max(a, b) < 3 + e^r$.

Proof: Let γ be the geodesic segment corresponding to V. Let g be the flowline corresponding to V. We can write $g = g_1|g_2$ where one of two things is true:

- g_1 is tiny and g_2 is empty.
- g_1 is small symmetric and g_2 is tiny.

We consider these two cases in turn.

Case 1: Let γ_1 be the geodesic segment corresponding to g_1, having endpoint (a, b, c). By the hyperbolic estimates, we know that $\eta_Y(\gamma)$ lies in the hyperbolic disk D_r in Π_Y centered at the origin. But the points in D_r all have X-coordinates at less than e^r. This gives $a < e^r$. We also have $b < \sqrt{5}$ once L is sufficiently large, by the Asymptotic Theorem. Hence $ab^2 < 5e^r$. We also see that $\max(a, b) < 3 + e^r$.

Case 2: Let γ_1 and γ_2 repectively be the geodesic segments corresponding to g_1 and g_2. Let r_j be the length of γ_j. Let (a_j, b_j) be the projection to η_Z of the far endpoint of γ_j.

Let γ_1' be the geodesic segment which is the first half of γ_1, in terms of length. So, γ_1' and γ_1 have the same initial endpoint (the origin) but γ_1' has length $r_1/2$. Let (a_1', b_1') be the far endpoint of $\eta_Z(\gamma_1')$. We make the following observations:

- By the same reasoning as in Case 1, we have $a_1' \leq e^{r_1/2}$. By the Doubling Lemma, $a_1 \leq 2e^{r_1/2}$.
- By the Asymptotic Theorem, $b_1 < 2 + \epsilon$ once L is large enough.
- By the same reasoning as in Case 1, except that we switch the roles of X and Y, we have $b_2 \leq e^{r_2}$.
- By symmetry and the Asymptotic Theorem, we have $a_2 < (2 + \epsilon)$ once L is sufficiently large.

Combining these observations, we have

$$a \leq (2 + \epsilon) + e^{r_1/2}, \qquad b \leq (2 + \epsilon) + e^{r_2}. \tag{32}$$

We get right away that $\max(a, b) < 3 + e^r$. The quantity $a^2 b$ is a finite sum of terms of the form $Ce^{\mu_1 r_1 + \nu_1 r_2}$, and in all cases $\mu_1 r_1 + \mu_2 r_2 \leq r$, and the sum of the coefficiients is

$$1 + 3(2 + \epsilon) + 3(2 + \epsilon)^2 + (2 + \epsilon)^3 < 28$$

The last inequality holds as long as we take $\epsilon < 1/100$. ♠

Once r is sufficiently large, the sphere S_r' consists entirely of small and perfect vectors either contained in the planes Π_X and Π_Y or else lying in loop level sets whose period is so large that Lemma 5.3 holds for them. Hence, for r sufficiently large, Lemma 5.3 shows that the projection of the positive sector of $\mathcal{S}_r$ lies in the region Ω_r defined by the following inequalities.

$$X, Y \in [0, 3 + e^r], \qquad \min(xy^2, yx^2) = 28e^r. \tag{33}$$

The region Ω_r is contained in a union of 2 overlapping regions which are interchanged by reflection in the main diagonal. One of the pieces is the

region underneath the graph $y = (28e^r/x)^{1/2}$ from starting from $x = c$ and ending at $x = 3 + e^r$. This region has area

$$\sqrt{2}8e^{r/2} \int_c^{3+e^r} \frac{dx}{\sqrt{x}} < 2 \times \sqrt{28} \times \sqrt{3 + e^r} < 12e^r,$$

once r is large. Hence Ω_r has area at most $24e^r$. Recalling that Ω_r contains the projection of the positive sector of $\mathcal{S}_r$, which is 1/4 of the whole sphere, we see that $A_X(\mathcal{S}_r) < 96e^r$ once r is sufficiently large.

5.3 The Yin Yang Curve

Given r, we define the *yin-yang curve* Y_r to be the set of points in S'_r where the differential $d(\eta_Z \circ E)$ is singular. For $r \leq \pi\sqrt{2}$ the the curve Y_r is connected. For $r > \pi\sqrt{2}$, the curve has 2 disjoint components interchanged by the map $(x, y, z) \to (y, x, -z)$.

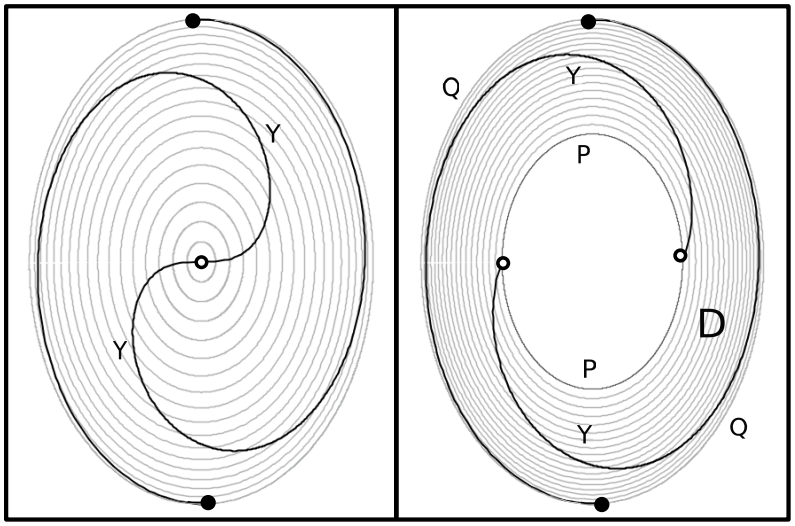

Figure 5.3: The yinyang curves for $r = \pi\sqrt{2}$ and $r = 5$.

Figure 5.3 shows the yin-yang curves for $r = \pi\sqrt{2}$ and for $r = 5$. In Figure 5.3, we are projecting the unit sphere in $\boldsymbol{R}^3$ onto the plane through the origin perpendicular to the vector $(1, -1, 0)$. The loop level sets all project to ellipses having aspect ratio $\sqrt{2}$. On the right side of Figure 5.3, the ellipse labeled P is the set of perfect vectors on S'_5. The ellipse labeled Q is the intersection of the positive sector of the unit sphere with the planes

$X = 0$ and $Y = 0$. Notice that $P_r \cup Y_r \cup Q_r$ divides S_r' into a union of 2 disks. The map $\eta_Z \circ \Pi$ is nonsingular on the interior of these disks and hence a local diffeomorphism. This is what is important for our Projection Estimate 4.

Referring to Figure 5.3, the union $Y_r \cup P_r \cup Q_r$ separates the positive sector of S_r' into 2 components. The map $(x, y, z) \to (x, y, -z)$ interchanges these components. Let D_r be either of these disks. Below, we will describe more clearly which of the two choices we take to be D_r. Figure 5.4 shows $\eta_Z \circ E(D_5)$. Essentially this is "half" of Figure 5.2.

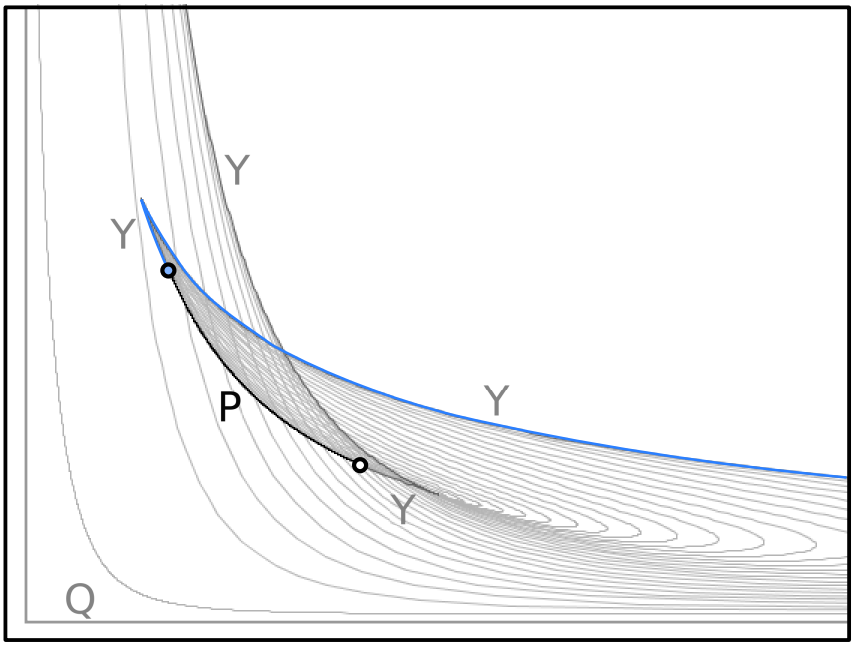

Figure 5.4: The image $\eta_Z \circ E(D_5)$.

The region labeled Y in Figure 5.3 is the image $\eta_Z \circ E(Y_r)$. We define $\Upsilon_r = \eta_Z \circ E(Y_r^*)$, where Y^* is the component of Y_r whose endpoint in Π_X is a point of the form $(x_r, y_r, 0)$ with $x_r > y_r$. We have drawn Υ_r in blue in Figure 5.4.

The map $L \to \Upsilon_r(L)$ is a smooth, and indeed real analytic, map. We say that a *cusp* of this map is a point where the map is not regular. So, away from the cusps, Υ_r is a smooth regular curve. Figure 5.3 suggests that Υ_5 just has a single cusp. We prove the following result in the next chapter.

Theorem 5.4 (Embedding) *For r sufficiently large, the curve Υ_r has a single cusp, and negative slope away from the cusp.*

5.4 The Multiplicity Bound

The image $\eta_Z \circ E(\partial D_r)$ is a piecewise analytic loop. We will show that this loop winds at most twice around any point in the plane that it does not contain. Referring to §2.3, we apply the Disk Lemma to $h = \eta_Z \circ E$ and $\Delta = D_r$. This tells us that $\eta_Z \circ E$ is at most 2-to-1 on D_r. But then $\eta_Z \circ E$ is at most 4-to-1 on the positive sector of S_r'. Hence η_Z is at most 4-to-1 on the positive sector of $\mathcal{S}_r$. Since the different sectors project into η_Z disjointly, we see that η_Z is at most 4-to-1 on all of $\mathcal{S}_r$. This establishes our estimate $N_Z(\mathcal{S}_r) = 4$.

Now we turn to the analysis of the image $\eta_Z \circ E(\partial D_r)$. Define

$$\Phi_r = E \circ \eta_Z(P_r). \tag{34}$$

Also, let $R(x, y, z) = (y, x, -z)$. The image $\eta_Z \circ E(\partial D_r)$ is invariant under R. It is the union of 5 analytic arcs:

- An arc of the X-axis connecting the origin to the endpoint of Υ_r.
- Υ_r.
- Φ_r.
- $R(\Upsilon_r)$.
- An arc of the Y-axis connecting an endpoint of $R(\Upsilon_r)$ to the origin.

Note that $\Upsilon_r \cup \Phi_r \cup R(\Upsilon_r)$ is a piecewise analytic arc that has its endpoints in the coordinate axes and otherwise lies in the positive quadrant. The Embedding Theorem says that Υ_r has negative slope, and is smooth and regular away from a single cusp. By symmetry, the Embedding Theorem also applies to $R(\Upsilon_r)$.

The cusps serve as natural vertices for our loop, so we make some new definitions which take the cusps into account. Let Φ_r^* denote the portion of $\Upsilon_r \cup \Phi_r \cup R(\Upsilon_r)$ that lies between the two cusps. Let $\Upsilon_r^* = \Upsilon_r - \Phi_r^*$. The loop $\eta_Z \circ E(\partial D_r)$ has the same 5-part description as above, with Υ_r^* and Φ_r^* used in place of Υ_r and Φ_r.

We will prove below that $\Upsilon_r^* \cup \Phi_r^*$ is embedded. By symmetry, $\Phi_r^* \cup R(\Upsilon_r^*)$ is also embedded. We will also prove that Υ_r^* crosses the main diagonal – the fixed point set of R – exactly once. This information forces the schematic picture of $\eta_Z \circ E(\partial D_r)$ shown in Figure 5.5.

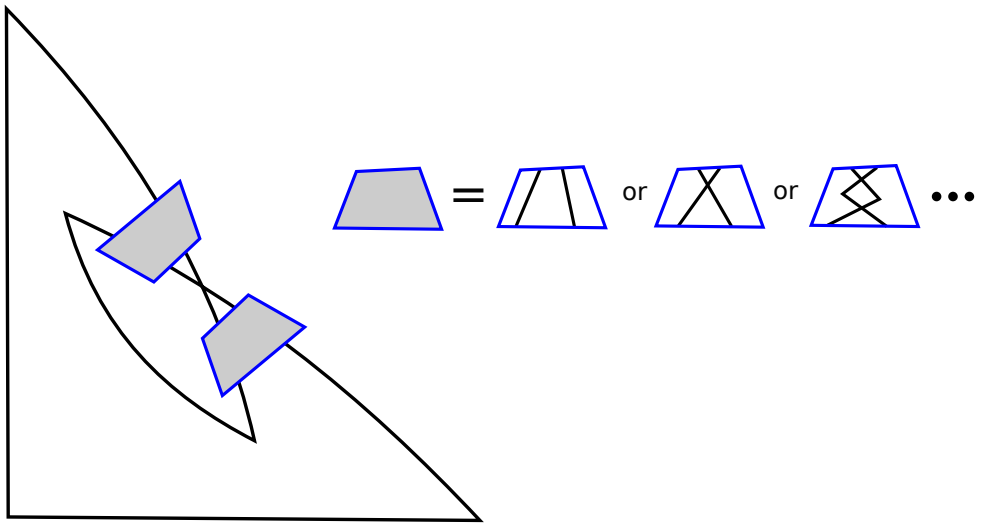

Figure 5.5: Schematic picture of $\eta_Z \circ E(D_5)$.

Numerically, it seems that the first option occurs, and that the two curves Υ_r^* and $R(\Upsilon_r^*)$ intersect exactly once. We did not want to take the trouble to establish this fact, given the already lengthy nature of the paper. In any case, the information above establishes the fact that $\eta_Z \circ E(\partial D_r)$ winds at most twice around any point in the plane that does not lie in its image. Applying the Disk Lemma to the map $h = \eta_Z \circ E$ and the disk D_r we see that h is at most 2-to-1 on D_r. But then h is at most 4-to-1 on $D_r \cup I(R_r) = S_r^{++}$. This completes the proof, modulo the properties of Υ_r^* and Φ_r^*. We now turn to the task of establishing the properties about the topology of this planar loop.

5.5 The Topology of the Boundary

Our proof uses the Embedding Theorem in a crucial way. Here is the main application.

Lemma 5.5 *For r sufficiently large, the curve $\Upsilon_r^* \cup \Phi_r^*$ is embedded.*

Proof: Our argument refers to Figure 5.6. Let γ be the portion of $\Upsilon_r^* \cup \Phi_r^*$ that lies above the (red) horizontal line L_1 through the cusp of $R(\Upsilon_r^*)$. This point is the endpoint of Φ_L^*.

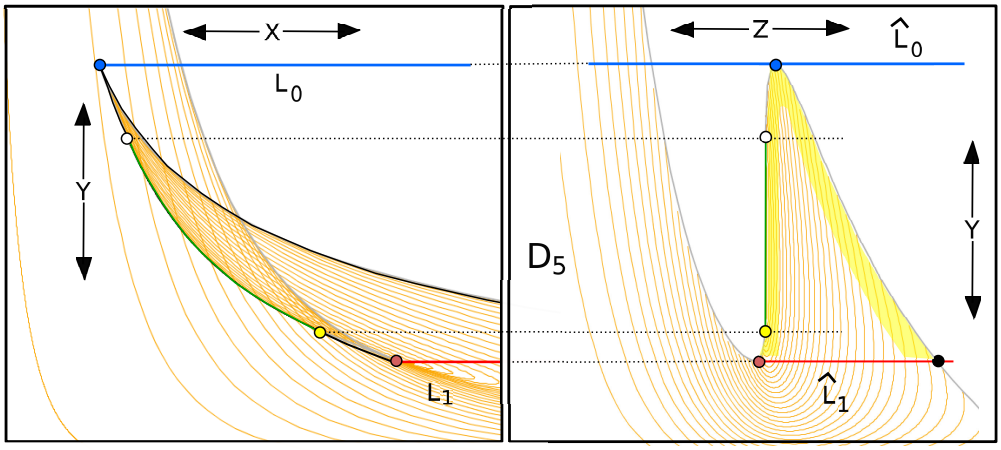

Figure 5.6: Projections of the relevant sets.

By the Embedding Theorem, γ has a single cusp, namely the cusp of Υ_r. Let γ_1 and γ_2 be the two arcs of γ on either side of this arc. These two arcs have negative slope and no cusps on them. Hence all of γ lies between the horizontal line L_0 through the cusp of Υ_r and the horizontal line through the cusp of $R(\Upsilon_r)$. These are the red and yellow horizontal lines on the left side of Figure 5.6.

Since γ_1 and γ_2 have negative slope and no cusps, they are each embedded. We just have to see that γ_1 cannot intersect γ_2. We will suppose that there is an intersection and derive a contradiction.

The portion of the Sol sphere $\mathcal{S}_r$ lying in the positive sector is the union of two disks, D_r and $R(D_r)$, where R is the isometry extending our reflection in the main diagonal of Π_Z, namely $I(x, y, z) = (y, x, -z)$. The common boundary of these disks contains a curve $\widehat{\gamma}$ which projects to γ. We have $\widehat{\gamma} = \widehat{\gamma}_1 \cup \widehat{\gamma}_2$. On the right side of Figure 5.6, one of these arcs connects the blue vertex to the black vertex, and the other one connects the blue vertex to the red vertex. The right side of Figure 5.6 shows the projection into the YZ plane.

Because no plane tangent to $\mathcal{S}_r$ at an interior point of $\widehat{\gamma}$ is vertical, each plane of the form $Y = \text{const.}$ intersects each of $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$ exactly once. In particular, this is true to the planes $\widehat{L}_0$ and $\widehat{L}_1$ which respectively project to L_0 and L_1 on the left side of Figure 5.6.

One of the two disks D_r or $R(D_r)$ has the property that it lies locally One of the two disks – say D_r – lies between $\widehat{L}_0$ and $\widehat{L}_1$ in a neighborhood

of $\widehat{\gamma}$. We are taking about the yellow highlighted region on the right side of Figure 5.6. The interior of D_r is transverse to the plane $\widehat{L}_1$ because η_Z is a local diffeomorphism on the interior of D_r. But this means that $D_r \cap \widehat{L}_1$ contains a smooth arc $\widehat{\beta}$ which connects the endpoint of γ_1 to the endpoint of γ_2. Let $\beta = \eta_Z(\widehat{\beta})$. We note the following

- β is contained in the line L_1.
- The endpoints of β coincide.
- The interior of β is a regular curve.

These properties are contradictory, because β would have to turn around in L_1 at an interior point, violating the regularity. This contradiction establishes the result that γ is embedded.

It remains to consider the portion of $\Upsilon_r \cup \Phi_r \cup R(\Upsilon_r^*)$ that lies below the horizontal line L_1 through the cusp of $R(\Upsilon_r^*)$. We label so that $\Phi_r \cup R(\Upsilon_r^*) \subset \gamma_2$. Given that γ_2 has negative slope we see that $\Phi_r \cup R(\Upsilon_r^*)$ lies entirely above L_1. But this means that the portion of γ_1 below L_1 is disjoint from γ_2. Finally, the portion of γ_1 below L_1 is disjoint from the portion of γ_1 above L_1 because γ_1 is regular and has negative slope. ♠

Define

$$\Upsilon_r^{**} = \Upsilon_r[r, 2r] - \Upsilon_r^*. \tag{35}$$

This is the subset of $\Upsilon_r[r, 2r]$ that occurs after the cusp. Given our result above, the only self-intersections on the curve $\eta_Z \circ D(\partial D_r)$ occur where Υ_r^{**} and $R(\Upsilon_r^{**})$. These are analytic arcs of negative slope, and they are permuted by the map R. Hence, the can only intersect finitely many times, and their intersection pattern must be as in Figure 5.6. This completes the proof of Projection Estimate 4, and hence the Volume Entropy Theorem, modulo the proof of the Embedding Theorem.

The rest of the paper is devoted to proving the Embedding Theorem.

6 The Embedding Theorem

6.1 The Isochronal Curves

Let $\underline{\Lambda}_L = (\underline{a}_L, \underline{b}_L)$, as in §3.5. Let E be the Riemannian exponential map and let η_Z be projection into the plane Π_Z. We have $\Upsilon_r = \eta_Z \circ E(Y_r^*)$, where Y_r^* is the relevant component of Y_r.

Lemma 6.1 $\Upsilon_r(L) = \underline{\Lambda}_L(r)$.

Proof: Recall that S_r' is the set of perfect vectors of length r contained in the positive sector of $\boldsymbol{R}^3$. Define the *positive side* of $S_r' \cap \Pi_Z$ to be those vectors of the form $(x, y, 0)$ with $x > y > 0$. By definition,

$$\Upsilon_r(L) = \eta_Z \circ E(Y_r^*), \tag{36}$$

where Y_r^* is the component of the yin yang curve Y_r which ends in the positive side. The vectors $V \in Y_r$ are characterized by the property that the differential $d(\eta_Z \circ E)$ is singular at points of Y_r.

The kernel of the projection map η_Z is spanned by the vector $(0, 0, 1)$. So, the differential $d(\eta_Z \circ E)$ is singular at V if and only if dE maps the tangent plane to S_r' at V to a plane which contains the vector $(1, 0, 0)$. But then $dE(N_V)$ is orthogonal to $(0, 0, 1)$. Here N_V is normal to S_r' at V. But this means that the third coordinate of $dE(N_V)$ is 0. Given the connection between the Haniltonian flow on S_r' and the geodesics, this situation happens if and only if the flowline associated to V ends in the plane Π_Z.

In short, Y_r consists of those small or perfect vectors of length r whose corresponding flowlines end in Π_Z. But these flowlines are then the initial halves of symmetric flowlines which wind at most twice around their loop level sets. The points in Y_r^* are the initial halves of symmetric flowlines whose midpoints lie on the positive side of S_r'. Moreover, these symmetric flowlines wind at most twice around their loop level sets and every amount of winding, so to speak, from 0 times to 2 times, is achieved. So, by definition $\Upsilon_r = \underline{\Lambda}(r)$. ♠

We call Υ_r an *isochronal curve* because it computes all the solutions to the differential equation at the fixed time r. Figure 6.1 shows part of Υ_5. The blue curves are the various curves $\underline{\Lambda}_L[0, L]$.

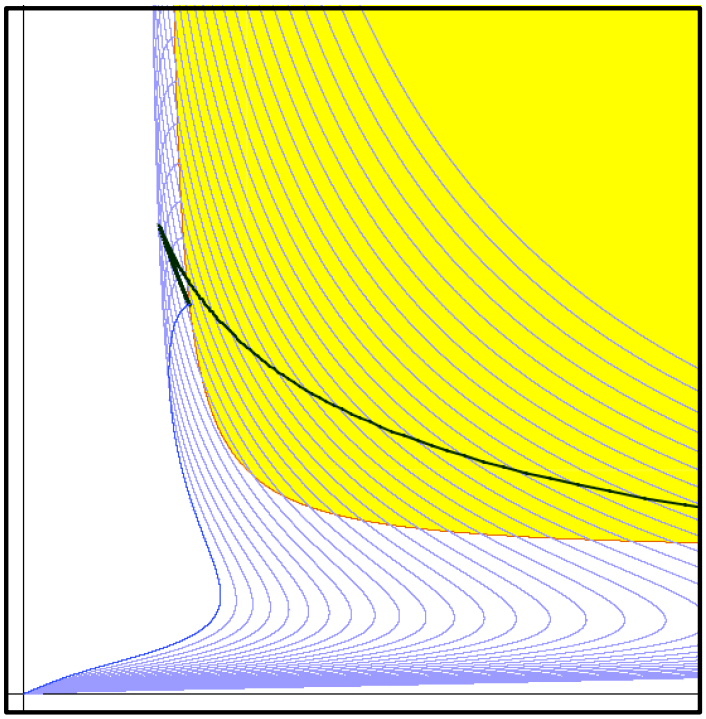

Figure 6.1: The $\underline{\Lambda}_L$ curves and the initial part of Υ_5.

6.2 The Tail End

Lemma 6.2 *The curve $\Upsilon_r(2r, \infty)$ is smooth, regular, and embedded.*

Proof: The flowline corresponding to the point $\Upsilon_r(L)$ lies on the loop level set of period L and flows for time r. If $L > 2r$ then the flowline travels less than halfway around the loop level set. Thus, the flowline is the initial half of a small symmetric arc. Let $S \subset \boldsymbol{R}^3$ denote the set of vectors corresponding to small symmetric flowlines. The map E is a diffeomorphism on S, because S consists entirely of small vectors. This is part of the Cut Locus Theorem from [**CS**]. By the Doubling Lemma,

$$\Upsilon_r(2r, \infty) = \frac{1}{2} E(C_r),$$

where $C_r \subset S$ is a smooth regular curve, obtained by dilating a suitably chosen arc of the yin yang curve by a factor of 2. Since E is a diffeomorphism on S, we see that $\frac{1}{2}E(C_r)$, is embedded. ♠

6.3 The Slope

Let x_L, etc. be the functions described at the end of §3.5. These functions satisfy the ODE

$$x' = -xz, \quad y' = yz, \quad z' = x^2 - y^2, \quad \underline{a}' = x + az, \quad \underline{b}' = y - bz, \quad (37)$$

with initial conditions

$$x_L(0) > y_L(0) > z_L(0) = \underline{a}_L(0) = \underline{b}_L(0) = 0, \quad (x_L(0), y_L(0), 0) \in \Theta_L. \quad (38)$$

Here Θ_L is the loop level set, in the positive sector, having period L. I am grateful to Matei Coiculescu for help with the following derivation.

Lemma 6.3 *Υ_r has negative slope away from the cusps.*

Proof: This proof is a calculation with the ODE. For any relevant function f, the notation $\dot{f}$ means $\partial f/\partial L$. We compute

$$(\dot{\underline{a}})' = \frac{\partial}{\partial L}\frac{\partial \underline{a}}{\partial t} = \frac{\partial}{\partial L}(x + z\underline{a}) = \dot{x} + \dot{\underline{a}}z + \dot{z}\underline{a}.$$

By the product rule

$$(x\dot{\underline{a}})' = x'\dot{\underline{a}} + x(\dot{\underline{a}})' = -zx\dot{\underline{a}} + x\dot{z} + xz\dot{\underline{a}} + \underline{a}x\dot{z} = x\dot{x} + \underline{a}x\dot{z}.$$

This calculation, and a similar one, show that

$$(x\dot{\underline{a}})' = x\dot{x} + \underline{a}x\dot{z}, \qquad (y\dot{\underline{b}})' = y\dot{y} - \underline{b}y\dot{z} \quad (39)$$

Since $x^2 + y^2 + z^2 \equiv 1$ we have $x\dot{x} + y\dot{y} + z\dot{z} = 0$. Adding the Equations in Equation 6.3 and using this relation, we find that.

$$(x\dot{\underline{a}} + y\dot{\underline{b}})' = x\dot{x} + y\dot{y} + (ax - by)\dot{z} = x\dot{x} + y\dot{y} + z\dot{z} = 0.$$

Hence $x\dot{\underline{a}} + y\dot{\underline{b}}$ is a constant function. Since $\underline{a}(0) = \underline{b}(0) = 0$ for all L we have $\dot{\underline{a}}(0) = 0$ and $\dot{\underline{b}}(0) = 0$. So, the constant in question is 0. Therefore

$$x\dot{\underline{a}} + y\dot{\underline{b}} = 0 \quad (40)$$

Because $\Upsilon_r(L) = \underline{\Lambda}_L(r)$, the velocity of Υ_r at L is

$$(\dot{\underline{a}}_L(r), \dot{\underline{b}}_L(r)).$$

This equals $-x_L(r)/y_L(r)$ by Equation 40. Since x, y are everywhere positive, the slope of Υ_r is negative whenever the velocity is nonzero. ♠

6.4 The End of Proof

In the next chapter we prove the following result.

Lemma 6.4 (Monotonicity) *If L is sufficiently large then there is some $t_L \in (L-1, L)$ such that the function $\underline{\dot{b}}_L$ is negative on $[L/2, t_L)$ and positive on $(t_L, L]$. Moreover, the function $L \to t_L$ is monotone increasing.*

Corollary 6.5 *$\Upsilon_r[r, 2r]$ has exactly one cusp.*

Proof: The curve Υ_r has a cusp at L if and only if its velocity

$$(\underline{\dot{a}}_L(r), \underline{\dot{b}}_L(r))$$

vanishes. By Lemma 40, one coordinate of the velocity vanishes if and only if the other one does. So, Υ_r has a cusp at L if and only if $\underline{\dot{b}}_L(r) = 0$. Suppose then that $\Upsilon_r[r, 2r]$ has more than one cusp. Then there are at least two pairs (r, L_1) and (r, L_2) such that $\underline{\dot{b}}_{L_1}(r) = 0$ and $\underline{\dot{b}}_{L_2}(r) = 0$. This means that $r = t_{L_1} = t_{L_2}$. But this contradicts the Monotonicity Lemma. Hence Υ_r has at most one cusp.

Since the map $L \to t_L$ is unbounded, each sufficiently large r lies in its image. But this means that for sufficiently large r there is some $L \in (r, r+1)$ such that $r = t_L$. This means that Υ_r has a cusp at L. Hence, once r is sufficiently large, Υ_r has exactly one cusp. ♠

This completes the proof of the Embedding Theorem, but there is one more remark we want to make. The cusp of Υ_r occurs at some $L \in (r, r+1)$. This explains why, in Figure 6.1, the cusp appears all the way to the left, near the end of Υ_5.

The next two chapters are devoted to the proof of the Monotonicity Lemma.

7 The Vanishing Point

7.1 Auxiliary Functions

In this chapter we will prove the first half of the Monotonicity Lemma. That is, we will show that once L is sufficiently large there is a point $t_L \in [L-1, L]$ such $\dot{b}_L$ vanishes at t_0, and this is the only vanishing point. We sometimes set $\dot{f} = \partial f/\partial L$ and $f' = \partial f/\partial t$. We introduce the following functions.

$$Z = \dot{z}, \quad X = \dot{x}/x, \quad Y = \dot{y}/y, \quad B = \dot{b}/b, \tag{41}$$

Since $a, b, x, y > 0$ these functions respectively have the same signs as $\dot{z}, \dot{x}, \dot{y}, \dot{b}$.

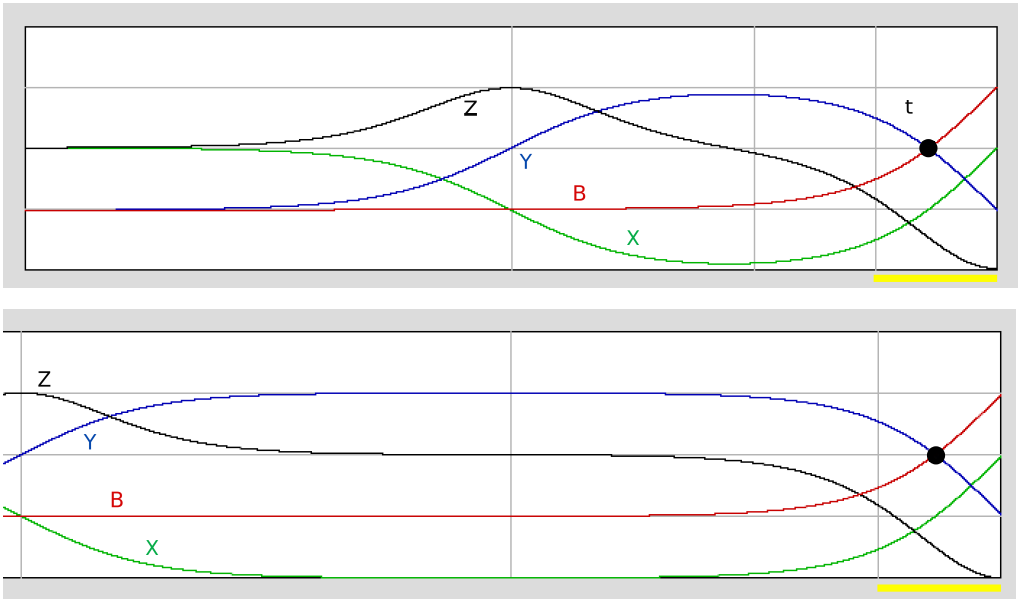

Figure 7.1: Numerical plots

The top half of Figure 7.1 shows numerical plots of these functions at $L = 8$. The bounding box in the picture is $[0, 8] \times [-1, 1]$. The bottom half shows the plot of the functions at the value $L = 16$. This time we are just showing the right half of the plot. Notice that in the interval $[L-1, L]$ the plots line up very nicely. The black dot in both cases is the point $(t_L, 0)$. One half of the Monotonicity Lemma establishes that t_L is uniquely defined. The second half shows that t_L increases monotonically. The intuition behind the second half of the result is that the pictures in $[L-1, L]$ stabilize, so that $t_L = L - s_L$ where $\partial s_L/\partial L$ is approximately 0 for large L.

7.2 Differentiation Formulas

In this section we derive the following formulas.

1. $X' = -Z$
2. $Y' = +Z$
3. $Z' = -4y^2Y - 2zZ.$
4. $B' = \frac{y}{\underline{b}}(Y - B) - Z.$
5. $Z'' = (-2 + 6z^2)Z.$

(1) From the relation $x' = -xz$ we get $(\dot{x})' = -\dot{x}z - \dot{z}x$. Then we get:

$$X' = \left(\frac{\dot{x}}{x}\right)' = \frac{-\dot{x}z - \dot{z}x}{x} - \frac{\dot{x}x'}{x^2} = -\dot{z} - \frac{\dot{x}z}{x} + \frac{\dot{x}z}{x} = -\dot{z} = -Z$$

(2) From the relation $y' = yz$ we get $(\dot{y})' = \dot{y}z + \dot{z}y$. Then we get:

$$Y' = \left(\frac{\dot{y}}{y}\right)' = \frac{\dot{y}z + \dot{z}y}{y} - \frac{\dot{y}y'}{y^2} = \dot{z} + \frac{\dot{y}z}{y} - \frac{\dot{y}z}{y} = \dot{z} = Z$$

(3) From the relations $z' = x^2 - y^2$ and $x^2 = 1 - y^2 - z^2$ we get

$$Z' = 2x\dot{x} + 2y\dot{y} = 2x^2X - 2y^2Y, \qquad 2x^2X = -2y^2Y - 2zZ.$$

Substitute the second relation into the first to get the formula above.

(4) Note that $\dot{\underline{b}}/\underline{b} = \dot{b}/b$ because $\underline{b} = b/2$. So, we work with b for ease of notation. From the relation $b' = 2y - bz$ we get $(\dot{b})' = 2\dot{y} - \dot{z}b - \dot{z}b$. Then:

$$B' = \left(\frac{\dot{b}}{b}\right)' = \frac{2\dot{y} - \dot{z}b - \dot{b}z}{b} - \frac{\dot{b}b'}{b^2} = \frac{2\dot{y} - \dot{z}b - \dot{b}z}{b} - \frac{2\dot{b}y - \dot{b}bz}{b^2} =$$

$$\frac{2\dot{y}}{b} - \frac{2\dot{b}y}{b^2} - \dot{z} = \frac{2y}{b}(Y - B) - Z = \frac{y}{\underline{b}}(Y - B) - Z.$$

(5) We first work out that

$$z'' = -2z + 2z^3. \tag{42}$$

We then compute

$$Z'' = \frac{\partial z''}{\partial L} = \frac{\partial(-2z + 2z^3)}{\partial L} = (-2 + 6z^2)Z.$$

7.3 The Formula for B

Here is a formula for B in terms of the other quantities. Matei Coiculescu found this formula and our derivation follows his ideas.

$$B = \frac{x\underline{a}}{2y\underline{b}}X - \frac{1}{2}Y - \frac{1}{2y\underline{b}}Z. \tag{43}$$

We will work with a and b rather than $\underline{a}$ and $\underline{b}$ until the end. We have

$$(ax - by)' = 2x^2 - 2y^2 = 2z', \qquad a(0) = b(0) = 0.$$

Integrating, we get

$$ax - by = 2z. \tag{44}$$

Given that $a = 2\underline{a}$ and $b = 2\underline{b}$, Equation 40 from the previous chapter is equivalent to:

$$x\dot{a} + y\dot{b} = 0. \tag{45}$$

Differentiating Equation 44 with respect to L, we have

$$x\dot{a} + a\dot{x} - y\dot{b} - b\dot{y} = 2\dot{z}. \tag{46}$$

Subtracting Equation 46 from Equation 45 we get

$$2y\dot{b} - a\dot{x} + b\dot{y} = -2\dot{z}. \tag{47}$$

Rearranging this, we get

$$\dot{b} = \frac{a\dot{x} - b\dot{y} - 2\dot{z}}{2y}.$$

When we make the substitutions

$$\underline{a} = a/2, \quad \underline{b} = b/2, \quad X = \dot{x}/x, \quad Y = \dot{y}/y, \quad Z = \dot{z},$$

we get Equation 43.

There is a similar formula for the function A, but we don't need it.

7.4 Asymptotics

If f and g are functions of L, we write $f \sim g$ if $f/g \to 1$ as $L \to \infty$. In this section we prove the following results.

$$(X_L(L/2), Y_L(L/2), Z_L(L/2), B_L(L/2)) \sim (-1/2, 0, 1/2, -1/2),$$

$$(X_L(L), Y_L(L), Z_L(L), B_L(L)) \sim (0, -1/2, -1, 1/2). \tag{48}$$

These various features are already apparent in Figure 7.1. Here we have written e.g. X_L in place of X to explicitly indicate how the quantity depends on L.

Lemma 7.1 $Y_L(0) = Y_L(L)$.

Proof: The limits we take for $Y_L(0)$ and $Y_L(L)$ respectively are

$$Y_L(0) = \lim_{\epsilon \to 0} \frac{y_{L+\epsilon}(0) - y_L(0)}{\epsilon y_L(0)}, \qquad Y_L(L) = \lim_{\epsilon \to 0} \frac{y_{L+\epsilon}(L) - y_L(L)}{\epsilon y_L(L)}. \tag{49}$$

Note also that $y_L(L) = y_L(0)$ and $y_{L+\epsilon}(L) = y_{L+\epsilon}(\epsilon)$. Hence

$$Y_L(L) = \lim_{\epsilon \to 0} \frac{y_{L+\epsilon}(\epsilon) - y_L(0)}{\epsilon y_L(0)}. \tag{50}$$

But the map $t \to Y_{L+\epsilon}(t)$ has a local minimum at $t = 0$ and so $y_{L+\epsilon}(\epsilon) = y_{L+\epsilon}(0) + O(\epsilon^2)$. Hence, the limit in Equation 50 equals $Y_L(0)$. ♠

Lemma 7.2 $X_L(L/2) = Y_L(0)$.

Proof: We have the relations

$$x_L(L/2) = y_L(0), \qquad x_{L+\epsilon}(L/2) = y_{L+\epsilon}(\epsilon/2) = y_{L+\epsilon}(0) + O(\epsilon^2).$$

The last equality comes from the fact that the function $t \to y_{L+\epsilon}(t)$ has its local minimum at 0. From these relations we have

$$\dot{x}_L(L/2) = \lim_{\epsilon \to 0} \frac{x_{L+\epsilon}(L/2) - x_L(L/2)}{\epsilon} = \lim_{\epsilon \to 0} \frac{y_{L+\epsilon}(0) + O(\epsilon^2) - y_L(0)}{\epsilon} = \dot{y}_L(0).$$

The lemma follows from this last equation and from $x_L(L/2) = y_L(0)$. ♠

Lemma 7.3 $Y_L(0) \sim -1/2$

Proof: We set $y = y_L(0)$. Using Equation 17 we get an explicit formula for L in terms of y. It is given by Equation 10. By the Inverse Function Theorem we have

$$Y_L(0) = \frac{1}{y \times \frac{\partial L}{\partial y}(y)}. \tag{51}$$

By Lemma 2.3, we get $Y_L(0) \sim -1/2$. ♠

Lemma 7.4 $X_L(0) \sim 0$.

Proof: We differentiate $x^2 + y^2 + z^2 = 1$ and use $z_L(0) = 0$ to get

$$x_L^2(0)X_L(0) + y_L^2(0)Y_L(0) = 0. \tag{52}$$

We see that $X_L(0) \sim 0$ because $x_L(0) \sim 1$ and $y_L(0) \sim 0$ and $Y_L(0) \sim -1/2$ (a finite number). ♠

Lemma 7.5 $Z_L(L) \sim -1$ *and* $Z_L(L/2) \sim 1/2$.

Proof: We have

$$z_L(L/2) = z_(L) = 0, \qquad z_{L+\epsilon}(L/2) = z_{L+\epsilon}(\epsilon/2), \qquad z_{L+\epsilon}(L) = -z_{L+\epsilon}(\epsilon).$$

Hence

$$-Z_L(L) = \lim_{\epsilon \to 0} \frac{z_{L+\epsilon}(\epsilon)}{\epsilon} = \lim_{\epsilon \to 0} \frac{z_{L+\epsilon}(\epsilon) - z_L(\epsilon)}{\epsilon} + \lim_{\epsilon \to 0} \frac{z_L(\epsilon)}{\epsilon}.$$

The second limit on the right is just $z_L'(0)$. The first limit is 0, because

$$z_L(\epsilon) = z_L'(0) + O(\epsilon^2), \qquad z_{L+\epsilon}(\epsilon) = z_{L+\epsilon}'(0)\epsilon + O(\epsilon)^2 = (z_L'(0))\epsilon + O(\epsilon)^2.$$

Hence

$$Z_L(L) = -z_L'(0) = y_L^2(0) - x_L^2(0) \sim -1.$$

The proof for $Z_L(L/2)$ is similar, and indeed $Z_L(L/2) = -(1/2)Z_L(L)$. ♠

The rest of the relations for X_L and Y_L follow from the ones we have established above, from Lemmas 7.1 and 7.2, and the fact that $X_L + Y_L$ is a constant function.

Lemma 7.6 *$B_L(L) \sim 1/2$ and $B_L(L/2) \sim -1/2$.*

Proof: Note that

$$(2\underline{a}_L(L/2), 2\underline{b}_L(L/2), x_L(L/2), y_L(L/2)) = (\underline{b}_L(L), \underline{a}_L(L), y_L(L), x_L(L)).$$

Hence, by the Reciprocity Lemma and the Asymptotic Theorem,

$$2x_L(L/2)\underline{a}_L(L/2) = 2y_L(L/2)\underline{b}_L(L/2) = x_L(L)\underline{a}_L(L) = y_L(L)\underline{b}_L(L) \sim 2.$$

Equation 43 now gives us

$$B_L(L) = \frac{1}{2}X_L(L) - \frac{1}{2}Y_L(L) - \frac{1}{2y_L(L)\underline{b}_L(L)}Z_L(L) \sim$$

$$\frac{1}{2}X_L(L) - \frac{1}{2}Y_L(L) - \frac{1}{4}Z_L(L) \sim 0 + (-1/2)(-1/2) + (-1/4)(-1) = 1/2. \quad (53)$$

The proof for $B_L(L/2)$ is similar. This completes the proof. ♠

7.5 Variation of the Z Coordinate

Lemma 7.7 *Let $\pi\sqrt{2} < L < M$. Then $z_M = z_L$ at most once on $(0, L)$.*

Proof: We already mentioned above that $z'' = -2z + 2z^3$. Let z_L and z_M be two solutions to this differential equation. Consider the ratio $\phi = z_M/z_L$. This quantity is positive on $(0, L/2]$. By L'hopital's rule we can continuously extend ϕ to 0, and we have $\phi(0) > 1$.

We have

$$\frac{d\phi}{dt} = \frac{W}{z_L^2}, \qquad W = z_L z_M' - z_M z_L', \quad (54)$$

If $z_L \neq 0$ then W and ϕ' have the same sign. We compute

$$W' = z_L z_M'' - z_M z_L'' = 2z_L z_M (z_M^2 - z_L^2). \quad (55)$$

Suppose there is some smallest time $t \in (0, L/2]$ where $z_L(t) = z_M(t)$. Note that $W' \neq 0$ on $(0, t)$. Also, $W' > 0$ on some small interval $(0, \epsilon)$ because $z_M'(0) > z_L'(0)$. Hence $W' > 0$ on $(0, t)$. Hence ϕ is increasing on $(0, t)$. In particular, $\phi(t) > 1$. Hence $z_M(t) > z_L(t)$, a contradiction.

If there is no time $t \in (L/2, L)$ where $z_M(t) = z_L(t)$, then we are done. Otherwise, let t_0 be the smallest such time. We have $z_L(t_0) = z_M(t_0) < 0$. Since $z_L(t_0 - \epsilon) < z_M(t_0 - \epsilon)$ for small $\epsilon > 0$, we have $z'_M(t_0) \leq z'_L(t_0) < 0$. Since these two functions are solutions of a second order ODE, namely Equation 42, they cannot have the same initial conditions at t_0. Hence $z'_M(t_0) < z'_L(t_0) < 0$.

Let $\zeta_L = -z_L$ and $\zeta_M = -z_M$. We consider these functions on the interval (t_0, L). These are solutions of the same differential equation, with initial conditions $\zeta_L(t_0) = \zeta_M(t_0)$ and $0 < \zeta'_L(t_0) < \zeta'_L(t_0)$. The same argument as above, applied to ζ_L and ζ_M, shows $z_L(t) > z_M(t)$ for $t \in (t_0, L]$. ♠

Lemma 7.8 (Z variation) *Z_L changes sign at most once on $[0, L]$ and $Z_L \geq 0$ on $[0, L/2]$.*

Proof: For the first statement, suppose there are 3 points t_1, t_2, t_3 where $Z_L(t_i)$ for $i = 1, 2, 3$ alternates sign. But then for ϵ sufficiently small, the difference $z_{L+\epsilon}(t_i) - z_L(t_i)$ also alternates sign for $i = 1, 2, 3$. This contradicts Lemma 7.7. So, there at most one $t_0 \in [0, L]$ where Z_L changes sign. The second statment follows from the analysis in Lemma 7.7, which showed that $z_M(t) > z_L(t)$ when $L < M$ and $t \in (0, L/2)$. ♠

7.6 Variation of the Y Coordinate

Lemma 7.9 (Y Variation) *Let $\delta_0 = 1/7$. If L is sufficiently large then:*

1. $|Y_L|, |Z_L| < 5$ *on* $[L-1, L]$.
2. $Y_L(L) < -\delta_0$.
3. $Y'_L < -\delta_0$ *on* $[L-1, L]$.
4. $Y_L(L-1) > \delta_0$.
5. $Y_L > 0$ *on* $(L/2, L-1]$.

Proof of Statement 1: Let $\phi_L(t) = -Z_L(L-1)$. This function satisfies the differential equation

$$\phi_L(0) = -Z_L(L) \sim 1, \quad \phi'_L(0) = Z'_L(L) \sim 0, \quad \phi''_L(t) = (-2 + 6z_L^2(t))\phi_L(t).$$

For the last equality we used the fact that $z_L^2(L-t) = z_L^2(t)$. The solutions to this equation converge in the C^∞ sense to the solutions of the equation

$$\phi(0) = 1, \qquad \phi'(0) = 0, \qquad \phi'' = (-2 + 6z^2)\phi. \tag{56}$$

Here z satisfies $z'' = -2z + 2z^3$ with initial conditions $z(0) = 0$ and $z'(0) = 1$. Since $\phi'' \in [-2, 4]\phi$, and since $\cos(t\sqrt{2}) > 0$ on $(0, 1]$, we have

$$\cos(t\sqrt{2}) \le \phi(t) \le \cosh(2t). \tag{57}$$

Since $\phi_L \to \phi$ we see that $|\phi_L| < 2$ once L is large. Hence $|Z_L| < 4$ on $[L-1, 1]$. Since $Y_L(L) \sim -1/2$ we have $|Y_L| < 5$ on $[L-1, L]$. ♠

Proof of Statement 2: This follows from the fact that $Y_L(L) \sim -1/2$.

Proof of Statement 3: Let ϕ_L and ϕ be the functions from the previous lemma. For $t \in [0, 1]$ we have (because ϕ is monotone decreasing)

$$Y_L'(L-t) = Z_L(L-t) = -\phi_L(t) < \epsilon - \phi(t) \le$$
$$\epsilon - \phi(1) \le \epsilon - \cos(\sqrt{2}) < -1/7 \tag{58}$$

once ϵ is sufficiently small. We can arrange this by taking L sufficiently large. ♠

Proof Statement 4: To estimate $Y_L(L-1)$ we note that

$$Y_L(L-1) = Y_L(L) - \int_{L-1}^{L} Z_L \, dt = Y_L(L) + \int_0^1 \phi_L \, dt >$$
$$-\epsilon - 1/2 + \int_0^1 \phi \, dt = -\epsilon - 1/2 + \frac{\sin(\sqrt{2})}{\sqrt{2}} > 1/7, \tag{59}$$

once ϵ is sufficiently small. ♠

Proof Statement 5: Suppose this is false. Since $Y_L(L/2) \ge 0$ and $Y_L'(L/2) = Z_L(L/2) > 0$ we see that Y_L is somewhere positive on $(L/2, L-1]$. Also, $Y_L(L-1) > 0$ and $Y_L(L) < 0$. If $Y_L = 0$ somewhere else on $(L/2, L-1]$ then Y_L switches signs at least 3 times. But then $Z_L = Y_L'$ switches sign at least twice. This contradicts the Z Variation Corollary. ♠

7.7 Variation of the B Coordinate

Lemma 7.10 (B variation) *As $L \to \infty$, the function $Y_L + B_L$ converges to 0 in the C^1 sense.*

Proof: We sometimes suppress the dependence on L in our notation. By Equation 48 we have $|Y(L) + B(L)| < \epsilon/2$ if L is sufficiently large. To finish the proof, it suffices to show that $|B' + Y'| < \epsilon/2$ on $[L-1, 1]$ for large L. Combining our derivative formulas with the preceding two results, we have

$$|B' + Y'| \leq \left|\frac{y}{\underline{b}}\right| \left(\max_{[L-1,L]} |Y| + |B| \right) \leq \frac{10y}{b} < 10y.$$

For the last inequality, we note that $b(L/2) > 1$ when L is large and $b' > 0$ on $[L/2, L]$. Hence $b > 1$ on $[L-1, 1]$. As $L \to \infty$ the maximum value of y on $[L-1, L]$ tends to 0. ♠

7.8 Uniqueness of the Vanishing Point

In this section we prove the first half of the Monotoniticy Theorem. That is, we show that $\dot{\underline{b}}_L$ vanishes exactly once on $(L/2, L)$. This is the same saying that B_L vanishes exactly once on $(L/2, L)$. Here is the crucial step.

Lemma 7.11 *B_L vanishes exactly once in $[L-1, L]$, at an interior point t_L, and $B_L(L-1) < 0$.*

Proof: Combining the Y Variation Lemma and the B Variation Lemma, we get that $B_L(L-1) < 0$ and $B_L(L) > 0$ and $\dot{B}_L > 0$ on $[L-1, L]$. ♠

Let $t_1 \in [L/2, L-1]$ be the smallest value such that $Z \leq 0$ on $[t_1, L-1]$.

Lemma 7.12 *$B_L < 0$ on $[t_1, L-1]$.*

Proof: Since $Z_L(L-1) < 0$ we know that $t_1 \in [L/2, L-1)$. We introduce the function $\phi(t) = -B_L(L-1-t)$. We have $\phi(0) > 0$ and $\phi'(t) = B_L'(L-1-t)$. Hence

$$\frac{\partial \phi}{\partial t} = \frac{y_L(L-1-t)}{\underline{b}_L(L-1-t)} \times \Big(Y_L(L-1-t) + \phi(t) \Big) - 2Z_L(L-1-t).$$

By the Y Variation Lemma and the definition of t_1 we have $\phi' = \mu_1\phi + \mu_2$ where μ_1 and μ_2 are non-negative functions on $[0, L-1-t_1]$. Since $\phi(0) > 0$ we have $\phi > 0$ on $[0, L-1-t_1]$. Hence $B_L < 0$ on $[t_1, L-1]$. ♠

If $t_1 = L/2$ this next lemma is vacuous.

Lemma 7.13 *$B_L < 0$ on $(L/2, t_1)$.*

Proof: Here is Equation 43 again:

$$B = \frac{x\underline{a}}{2y\underline{b}}X - \frac{1}{2}Y - \frac{1}{2y\underline{b}}Z.$$

Our result here follows from Equation 43 and these inequalities on $[L/2, t_1)$:

- The quantities $x_L, y_L, \underline{a}_L, \underline{a}_L$ are all positive.
- Since $X_L(L/2) \le 0$ and $X_L' = -Z_L \le 0$ on $(L/2, t_1)$, we have $X_L \le 0$.
- By the Y Variation Lemma, $Y_L > 0$.
- Since Z_L changes sign only at t_1, and $Z_L(L) < 0$, we have $Z_L \ge 0$.

This completes the proof. ♠

8 Monotonicity of the Vanishing Point

8.1 The Proof Modulo Asymptotics

We know there is a unique $t_L \in (L-1, L)$ where $\dot{\underline{b}}_L$ vanishes. In this chapter we show that t_L is monotone increasing. Let

$$\beta_L(t) = -B_L(L-t). \tag{60}$$

There is a unique $s_L \in (0,1)$ such that $\beta_L(s_L) = 0$. In fact, $t_L = L - s_L$. Since $\dot{t}_L = 1 - \dot{s}_L$, it suffices to show $|\dot{s}_L| < 1$ for large L. We show $\dot{s}_L \sim 0$. That is, $\lim_{L\to\infty}(\dot{s}_L) = 0$. Define

$$\delta(t) = \frac{y(L-t)}{\underline{b}(L-t)}, \qquad \phi(t) = Y(L-t), \qquad \zeta(t) = Z(L-t). \tag{61}$$

Since $\beta'(t) = B'(L-t)$, we have

$$\beta' = \delta(\phi - \beta) - \zeta, \qquad (\dot{\beta})' = \dot{\delta}(\phi - \beta) + \delta(\dot{\phi} - \dot{\beta}) - \dot{\zeta}. \tag{62}$$

The second equation comes from differentiating the first with respect to L.

Let $\|f\|$ denote the sup of f on $[0,1]$. We will establish the following estimates below.

$$\|\delta\|, \|\dot{\delta}\|, \|\dot{\phi}\|, \|\dot{\zeta}\|, \dot{\beta}(0) \sim 0. \tag{63}$$

We also know from the Y Variation Lemma that $\|\phi\|, \|\zeta\| < 5$. Hence

$$(\dot{\beta})' = \epsilon_1 \dot{\beta} + \epsilon_2, \tag{64}$$

where ϵ_1 and ϵ_2 are functions such that $\|\epsilon_1\|, \|\epsilon_2\| \sim 0$. Given our initial condition $\dot{\beta}(0) \sim 0$, a standard comparison argument now says that $\|\dot{\beta}\| \sim 0$.

By definition

$$\beta_L(s_L) = 0. \tag{65}$$

Using implicit differentiation, we see that

$$|\dot{s}_L| = \left|\frac{\dot{\beta}(s_L)}{\beta'(s_L)}\right| < 8|\dot{\beta}(s_L)| \sim 0. \tag{66}$$

The last inequality comes from the fact that $|\beta'| > 1/8$ on $[L-1, L]$ once L is large enough. This proves the Monotonicity Lemma, modulo Equation 63.

8.2 The Asymptotics

We say that a function f of L is *tame* if $df/dL \sim 0$. If f is analytic (i.e. not contrived in an artificial way) and $f \sim$ const. one might expect f to be tame. We verify that this is the case for a number of quantities we have studied in the previous chapter. We also point out, when relevant, how the given quantity relates to the functions $\beta, \delta, \phi, \zeta$ introduced above.

Lemma 8.1 *$Y_L(L)$ is tame. Hence $\dot{\phi}(0) \sim 0$.*

Proof: Since $Y_L(0) = Y_L(L)$, it suffices to prove that $Y_L(0)$ is tame. Lemma 2.3 tells us that

$$\frac{d}{dy} Y_L(0) \sim 0, \tag{67}$$

where $y = y_L(0)$. But, since $Y_L(0)$ is asymptotic to a finite number, and $y_L(0) \sim 0$, we have $\frac{dy}{dL} \sim 0$. Therefore, *a fortiori* we have $\frac{d}{dL} Y_L(0) \sim 0$. ♠

Lemma 8.2 *$X_L(L)$ is tame.*

Proof: Since $X_L(0) = X_L(L)$, it suffices to prove that $X_L(0)$ is tame. Differentiating Equation 52 with respect to L, we get

$$2x_L^2(0)X_L^2(0) + x_L^2(0)\left(\frac{d}{dL}X_L(0)\right) + 2y_L^2(0)Y_L^2(0) + y_L^2\left(\frac{d}{dL}Y_L(0)\right) = 0. \tag{68}$$

Given that $x_L(0) \sim 1$ and $y_L(0) \sim 0$ and $Y_L(0) \sim -1/2$ and that $Y_L(0)$ is tame, we see that $X_L(0)$ is also tame. ♠

Lemma 8.3 *$Z_L(L)$ is tame. Hence $\dot{\zeta}(0) \sim 0$.*

Proof: We have

$$\frac{d}{dL} Z_L(L) = \frac{d}{dL}(y_L^2(0) - x_L(0)^2) = y_L^2(0)Y_L(0) - x_L^2(0)X_L(0) \sim 0.$$

The last equation comes from the fact that all quantities in the last expression are asymptotic to finite numbers and $y_L(0) \sim 0$ and $X_L(0) \sim 0$. ♠

Lemma 8.4 *$Z_L'(L)$ is tame. Hence $(\dot\zeta)'(0) \sim 0$.*

Proof: Since $z_L(L) = 0$ we have

$$Z_L'(L) = -4y_L^2(L)Y_L(L). \tag{69}$$

Differentiating Equation 69 with respect to L and using the product rule, as above, we see that $\frac{d}{dL}Z_L' \sim 0$. ♠

Lemma 8.5 *$y_L(L)\underline{b}_L(L)$ is tame.*

Proof: We begin by proving an estimate that will come in at the end of the proof. We claim that

$$\max_{[0,L/4]} |Y_L| < 1. \tag{70}$$

To see this, note that $Y_L' = Z_L$. We also know that $Z_L \geq 0$ on $[0, L/2]$. Hence Y_L is monotone increasing on $[0, L/2]$ and $Y_L(L/2) \sim 1/2$. This establishes Equation 70.

By Lemma 5.1, we see that

$$y_L(L)b_L(L) = 2z_L(L/4) + 2\phi_L, \qquad \phi_L = \int_0^{L/4} y_L^2 \, dt. \tag{71}$$

We deal with these terms one at a time. Referring to Equation 17 we have $z_L(L/4) = \sqrt{1 - 2\alpha_L^2}$. Here $\alpha_L \sim 0$. This leads to $\frac{d}{d\alpha}(Z_L(L/4)) \sim 0$. A calculation like the one done in Lemma 7.3 shows that $|dL_\alpha/d\alpha| \sim \infty$. Hence $d\alpha_L/dL \sim 0$. But then, by the chain rule, $\frac{d}{dL}z_L(L/4) \sim 0$.

We have

$$\frac{d\phi_L}{dL} = \int_0^{L/4} \frac{\partial}{\partial L}(y_L)^2 \, dt + \frac{1}{4}y_L(L/4). \tag{72}$$

Referring to Equation 17 we have $y_L(L/4) = \alpha$. So, the same argument as for $Z_L(L/4)$ now shows that $\frac{d}{dL}y_L(L/4) \sim 0$. Finally, we have

$$\int_0^{L/4} \frac{\partial}{\partial L}(y_L)^2 \, dt = 2\int_0^{L/4} y_L^2 Y_L \, dt \leq^* 2\int_0^{L/4} y_L^2 \sim 0.$$

The starred inequality comes from Equation 70. The final asymptotic result comes from the proof of Lemma 5.1. ♠

All of the asymptotic results we have obtained so far feed into one final one.

Lemma 8.6 *$B_L(L)$ is tame. Hence $\dot\beta(0) \sim 0$.*

Proof: As in Lemma 7.6 have

$$B_L(L) = X_L(L) - \frac{1}{2}Y_L(L) - \frac{1}{y_L(L)\underline{b}_L(L)}Z_L(L) \sim$$

$$X_L(L) - \frac{1}{2}Y_L(L) - \frac{1}{4}Z_L(L) \sim 0 + (-1/2)(-1/2) + (-1/4)(-1) = 1/2. \quad (73)$$

We know that

$$X_L(L), Y_L(L), Z_L(L), y_L(L)\underline{b}_L(L)$$

are all tame. Also, we know that

$$y_L(L)\underline{b}_L(L) \sim 2,$$

by the Asymptotic Theorem. Using all this information, and the product and quotient rules for differentiation, we see that $B_L(L)$ is tame. ♠

Lemma 8.7 *$\|Z\| \sim 0$.*

Proof: Our notation here is a bit funny. We mean to restrict our function Z_L to the interval $[0,1]$ and take its maximum. We have $Z(0) \sim 0$ and $|Z''| \le 4|Z]$. Since $Z > 0$ on $(0,1]$, the same kind of comparison argument used in the proof of the Y Variation Lemma now shows that $\max_{[0,1]} |Z| \le Z(0)\cosh(2) \sim 0$. ♠

Now we establish the remaining estimates from §8.1

Lemma 8.8 *$\|\delta\| \sim 0$ and $\|\dot\delta\| \sim 0$.*

Proof: The argument in the proof of the B Variation Lemma shows, incidentally, that $\|\delta\| \sim 0$. Using our derivative formulas, we have

$$\dot\delta(t) = (Y(L-t) - B(L-t))\delta(t).$$

Combining the Y Variation Lemma and the B Variation Lemma we see that $|B|, |Y| < 6$ on $[L-1, L]$ for large L. Hence $\|\dot\delta\| < 12\|\delta\| \sim 0$. ♠

Lemma 8.9 $\|\dot{\zeta}\| \sim 0$.

Proof: Define $\eta = \dot{\zeta}$. From the differential equation for Z'' and the fact that $z_{L-t} = z_t$ we get $\zeta''(t) = (-2 + 6z^2(t))\zeta$. Differentiating with respect to L and using the fact that the mixed partials commute, we see that

$$\dot{\zeta}''(t) = (-2 + 6z^2(t)\dot{\zeta}(t) + 12Z(t)z^2(t)\zeta(t).$$

The first term on the right lies in $[-4, 4]\dot{\zeta}(t)$. The second term is at most

$$12\|Z\| \max_{[L-1,L]} |Z| < 60\|Z\| \sim 0.$$

Here we have used Statement1 of the Y Variation Lemma and also the bound from Lemma 8.7. Putting these estimates together, we get

$$|\dot{\zeta}''| \leq 4|\dot{\zeta}(t)| + \epsilon, \tag{74}$$

where $\epsilon \sim 0$. We have already seen that $\dot{\zeta}(0) \sim 0$ and $(\dot{\zeta})'(0) \sim 0$. The same kind of comparison argument as above now give us the desired bound on $\dot{\zeta}$. ♠

Lemma 8.10 $\|\dot{\phi}\| \sim 0$.

Proof: We have $\phi' = -\zeta$. Differentiating with respect to L we get

$$(\dot{\phi})' = -\dot{\zeta}.$$

We also have $\dot{\phi}(0) \sim 0$. We now integrate the bound on $\|\dot{\zeta}\|$ to get the bound on $\|\dot{\phi}\|$. ♠

With these bounds, we complete the proof of the Monotonicity Lemma.

9 References

[**A**] V. I. Arnold, *Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits.* Ann. Inst. Fourier Grenoble, (1966).

[**AK**] V. I. Arnold and B. Khesin, *Topological Methods in Hydrodynamics*, Applied Mathematical Sciences, Volume 125, Springer (1998)

[**B**] N. Brady, *Sol Geometry Groups are not Asynchronously Automatic*, Proceedings of the L.M.S., 2016 vol 83, issue 1 pp 93-119

[**BB**] J. M. Borwein and P. B. Borwein, *Pi and the AGM*, Monographies et Études de la Société Mathématique du Canada, John Wiley and Sons, Toronto (1987)

[**BS**] A. Bölcskei and B. Szilágyi, *Frenet Formulas and Geodesics in Sol Geometry*, Beiträge Algebra Geom. 48, no. 2, 411-421, (2007).

[**BT**], A. V. Bolsinov and I. A. Taimanov, *Integrable geodesic flow with positive topological entropy*, Invent. Math. **140**, 639-650 (2000)

[**CMST**] R. Coulon, E. A. Matsumoto, H. Segerman, S. Trettel, *Noneuclidean virtual reality IV: Sol*, math arXiv 2002.00513 (2020)

[**CS**] M. P. Coiculescu and R. E. Schwartz, *The Spheres of Sol*, submitted preprint, 2020

[**EFW**] D. Fisher, A. Eskin, K. Whyte, *Coarse differentiation of quasi-isometries II: rigidity for Sol and Lamplighter groups*, Annals of Mathematics 176, no. 1 (2012) pp 221-260

[**G**], M. Grayson, *Geometry and Growth in Three Dimensions*, Ph.D. Thesis, Princeton University (1983).

[**K**] S. Kim, *The ideal boundary of the Sol group*, J. Math Kyoto Univ 45-2 (2005) pp 257-263

[**KN**] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry, Volume 2*, Wiley Classics Library, 1969.

[**LM**] R. López and M. I. Muntaenu, *Surfaces with constant curvature in Sol geometry*, Differential Geometry and its applications (2011)

[**S**] R. E. Schwartz, *Java Program for Sol*, download (in 2019) from
http://www.math.brown.edu/∼res/Java/SOL.tar

[**T**] M. Troyanov, *L'horizon de SOL*, Exposition. Math. 16, no. 5, 441-479, (1998).

[**Th**] W. P. Thurston, *The Geometry and Topology of Three Manifolds*, Princeton University Notes (1978). (See
http://library.msri.org/books/gt3m/PDF/Thurston-gt3m.pdf
for an updated online version.)

[**W**] S. Wolfram, *The Mathematica Book, 4th Edition*, Wolfram Media and Cambridge University Press (1999).