Statistical mechanics of billiard-like systems

T. Chumley,* S. Cook,† R. Feres‡

May 8, 2014

These are lecture notes for a short course to be given at the Houston Summer School on Dynamical Systems, at the University of Houston, May 14-22, 2014. Scott Cook, of Swarthmore College, Tim Chumley, of Iowa State University, and myself will be available throughout the week for discussions outside of the lectures.

1 Plan of lectures

1. Classical mechanical systems with boundary. Thursday, May 15, 9:30-10:30AM

   General framework of classical Newtonian mechanical systems in the Lagrangian formalism, with an emphasis on systems with elastic collisions (billiard systems). Of primary importance for what will follow is a description of natural (Hamiltonian flow-) invariant measures. We also cover a number of general ideas related to billiard dynamical systems.

2. Mechanics and probability. Monday, May 19, 10:45-11:45AM

   Probability theory comes into the picture in two ways. First, it is already present in the analysis of deterministic systems, in the context of ergodic theory. We show some of the ergodic theoretic consequences of the existence of finite invariant measures (Poincaré recurrence, Birkhoff's ergodic theory) for mechanical systems. Probability also comes in more directly, by regarding some of the dynamical variables as random variables with specified probability distribution. This will result in Markov chains with special properties that will serve as models of thermodynamical phenomena.

3. Models of thermal interaction. Tuesday, May 20, 1:15-2:15PM

   We discuss simple random mechanical models that allow us to think about thermal interaction in terms of billiard-Markov chains. We describe a very simple models of thermostats and study some of their properties, such as convergence to thermal equilibrium.

4. Diffusion and first steps into thermodynamics. Thursday, May 21, 9:30-10:30AM.

   Up to this point we will have mainly been concerned with "microscopic" phenomena. The transition to "macroscopic" phenomena (in the context of classical kinetic theory of "billiard" gases) requires passage to a diffusion limit. This relates to central limit theorems.

*Department of Mathematics, Iowa State University, Ames, IA 50011
†Department of Mathematics and Statistics, Swarthmore College, College Ave, Swarthmore PA 19081
‡Department of Mathematics, Washington University, Campus Box 1146, St. Louis, MO 63130

We describe diffusions in velocity space (Maxwell-Boltzmann diffusion, a concept we introduce in the lectures) as well as in position space. If time permits, we may very briefly explore ideas in non-equilibrium thermodynamics and stochastic thermodynamics.

2 Newtonian mechanical systems and geometry

We give a general definition of a mechanical system of Newtonian type. These are systems defined in terms of a Lagrangian function of the form: kinetic minus potential energy. We need a small amount of manifold theory to describe these systems in a clear and conceptual way. Nevertheless, the mechanical systems that will be the focus of our study later on (billiard-like systems) will require next to no acquaintance with differential geometry. In class we will discuss only a small amount of the extended overview given below.

2.1 Configuration manifolds and the Kinetic energy metric

It may be appropriate to begin by stating a general definition of what we are going to understand by a *mechanical system* and, more specifically, a *billiard-like* system (or a system of *billiard type*). This section will contain more differential geometry than strictly necessary for the rest of the lectures, but it may, we hope, provide some helpful perspective. Briefly, we use the term *Newtonian mechanical system* to refer to non-relativistic, classical (as opposed to quantum) systems whose Lagrangian function (to be introduced shortly) has the form of a kinetic energy function minus a potential function. In particular, we do not consider dissipative forces such as friction or viscosity. For billiard type systems, potential forces will be absent and the only type of interaction between the material parts will be through elastic collisions.

Definition 2.1 (System of masses and manifold of configurations)**.** The following data define a system of masses comprising the mechanical system and the space of possible configurations theses masses can assume in coordinate space $\mathbb{R}^n$. Typically, $n \leq 3$.

1. $\mathcal{B}$ denotes a Borel measurable space with a finite (positive) measure μ that specifies the mass distribution. The *total mass* is $m = \mu(\mathcal{B})$. We think of $\mathcal{B}$ as the *material body* in some fixed, or reference, configuration.

2. The *configuration manifold* M is a smooth manifold with *corners* as defined in [5]. (The boundary may be empty.) The dimension of M specifies the *degrees of freedom* of motions of the body.

3. The *position map* $\Phi : M \times \mathcal{B} \to \mathbb{R}^n$ assigns to each *material point* $b \in \mathcal{B}$ and $q \in M$ the position $\Phi(q, b)$ of b in the configuration q. We assume that $q \mapsto \Phi(q, b)$ is a smooth map for each b. Any of the equivalent notations $\Phi(q, b) = \Phi_q(b) = \Phi_b(q) = q(b)$ may be used.

The following simple example (see Figure 1) should help to clarify Definition 2.1. It consists of a pendulum with a circular bob of radius r rigidly attached to the end of a rod so that the distance from the pivot to the center of the bob is l. The pendulum pivot rotates freely (without friction) and slides freely along a horizontal rail. The sliding part of the system to which the

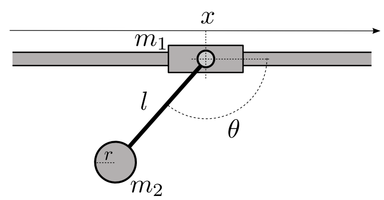

Figure 1: A disc of mass m is attached to the end of a rigid rod that swings freely about a pivot point. The pivot is mounted on a rail, along which it can slide without friction. The distance from the pivot to the center of the disc is l. The configuration manifold is topologically a cylinder parametrized by $(x,[\theta]) \in \mathbb{R}\times\mathbb{T}$, where $\mathbb{T}$ denotes the interval $[0,2\pi]$ with the endpoints 0 and 2π identified.

pivot is attached has mass m_1, assumed (with no loss of generality) to be concentrated at the pivot point. The rest of the mass of the system, m_2, is distributed uniformly over the pendulum bob. We let $\mathcal{B}_1 = \{0\}$ (the origin of $\mathbb{R}^2$) represent the pivot and $\mathcal{B}_2 = \left\{(b_1,b_2)\in\mathbb{R}^2 : b_1^2+b_2^2 \le r^2\right\}$ the bob, and write $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_1$, were the union should be understood to be disjoint. (Note that the various components of $\mathcal{B}$ are allowed to overlap. Strictly speaking, one should add an index to these components in order to be clear that one is dealing with a disjoint union, something we will not do for the sake of keeping the notation simple. Typically, $\Phi(q,\cdot)$ q will separate these overlapping components.) The mass distribution is then $d\mu(b) = (m_2/\pi r^2)\, db_1\, db_2$ for $b\in\mathcal{B}_2$ and $\mu(0) = m_1$ if $0\in\mathcal{B}_1$. The manifold of configurations is $M = \mathbb{R}\times\mathbb{T}$, where $\mathbb{T} = \mathbb{R}/2\pi\mathbb{Z}$. Equivalently, we may view $\mathbb{T}$ as the unit circle, so M is diffeomorphic to a cylinder. A point $q = (x,[\theta])$ represents a configuration in which the pendulum pivot is at position x along the rail and the pendulum arm is at an angle θ relative to the line of the rail. Here $[\theta]$ represents the equivalence class module 2π.

We may describe the position map for this example as follows. Let $e_1 = (1,0)^\dagger$, $e_2 = (0,1)^\dagger$ be the standard basis vectors in $\mathbb{R}^2$, regarded as column vectors. (The superscript '†' indicates matrix transpose.) Let R_θ denote the 2-by-2 matrix representing rotation counterclockwise by angle θ and $b = (b_1,b_2)^\dagger$. Then

$$\Phi((x,[\theta]),b) = \begin{cases} x & \text{if } b\in\mathcal{B}_1 \\ R_\theta(b+le_1)+xe_1 & \text{if } b\in\mathcal{B}_2 \end{cases} \tag{2.1}$$

2.2 THE EULER-LAGRANGE EQUATIONS

Definition 2.2 (Motions and states)**.** Given a mechanical system $(\mathcal{B},\mu,\Phi)$ with configuration manifold M, we define a *motion* of the system to be a path $\gamma : I \to M$ where I is some interval in $\mathbb{R}$. A (pure) *state* of the system is an element of the tangent bundle TM. We denote a state by the pair (q,v), where $q\in M$ and $v\in T_qM$. (Later on, a state will be understood more generally as a probability measure on M. If the measure is supported on a single point we say that it is a pure state.)

Given a state $\xi = (q,v)$, we introduce a vector field on $\mathcal{B}\subset\mathbb{R}^n$ as follows. Let $\gamma : (-a,a)\to M$

be a path *representing* (q,v). That is, a differentiable path such that $\gamma(0)=q$ and $\gamma'(0)=v$. Then the vector field, which we also denote by ξ, is

$$\xi(b) := \left.\frac{d}{dt}\right|_{t=0} \Phi(\gamma(t),b).$$

We think of the path parameter t as 'time' and call $b \mapsto \xi(b)$ the *velocity field* on $\mathcal{B}$ corresponding to the state $\xi \in TM$. If $\gamma(t)$ is given in terms of coordinates $(q_1,\ldots,q_n)$ on M, we typically denote time derivative of the coordinate functions by $\dot{q}_j$.

Definition 2.3 (Kinetic energy metric)**.** Given $(\mathcal{B},\mu,\Phi)$ with configuration manifold M, we define on TM a field of symmetric bilinear forms $q \mapsto \langle\cdot,\cdot\rangle_q$ as follows. For any two states $\xi=(q,u), \eta=(q,v)$ at $q \in M$ we define

$$\langle u,v\rangle_q := \int_{\mathcal{B}} \xi(q)\cdot\eta(b)\, d\mu(b)$$

where $\xi(q)\cdot\eta(b)$ indicates the standard inner product (dot-product) in $\mathbb{R}^n$. We always make the (very weak) assumption that the position map is *effective* in that $\langle\cdot,\cdot\rangle_q$ is a non-degenerate symmetric form at each q. In other words, it defines a Riemannian metric on M. The *kinetic energy* function is by definition the function $K : TM \to [0,\infty)$ given by

$$K(q,v) = \frac{1}{2}\|v\|_q^2.$$

Notice how the mass distribution given by the measure μ is incorporated into the norm $\|\cdot\|$.

As an example let us calculate the kinetic energy function for the system of Figure 1. The rotation matrix is $R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ and its time derivative for a motion described in coordinates by $(x(t),[\theta(t)])$ is the matrix $\dot\theta R_{\theta+\frac{\pi}{2}}$. Therefore, taking the derivative in t (at $t=0$) of $\Phi((x(t),[\theta(t)]),b)$ (see Equation 2.1) gives

$$\xi(b) = \begin{cases} \dot{x}e_1 & \text{if } b = 0 \in \mathcal{B}_1 \\ \dot\theta R_{\theta+\frac{\pi}{2}}(b+le_1)+\dot{x}e_1 & \text{if } b \in \mathcal{B}_2 \end{cases}$$

A simple calculation now gives the kinetic energy function evaluated at a tangent vector $v=(\dot{x},\dot\theta)$ of M, expressed in the coordinates $q=(x,\theta)$, as

$$K(q,v) = \frac{1}{2}\int_{\mathcal{B}} |\xi(b)|^2\, d\mu(b) = \frac{1}{2}m_1\dot{x}^2 + \frac{1}{2}m_2\left[\left(\frac{r^2}{2}+l^2\right)\dot\theta^2 + \dot{x}^2 - 2\dot{x}\dot\theta l\sin\theta\right].$$

The dynamics of the mechanical system is introduced through a choice of *Lagrangian* function on the phase space TM.

Definition 2.4 (Lagrangian and action functional)**.** The Lagrangian of a Newtonian mechanical system is the function $L : TM \to \mathbb{R}$ defined on the tangent bundle (the *phase space* of the system) TM and given by

$$L(q,v) := K(q,v) - V(q)$$

where $V(q)$ is a *potential energy* function on M. If $\gamma : [a,b] \to M$ is a differentiable path describing a motion of the system, the *action functional* evaluated on γ is

$$S[\gamma] := \int_a^b L(\gamma(t), \gamma'(t))\, dt.$$

According to Hamilton's principle of stationary action, a path $\gamma(t)$ describing an actual physical motion of the system must be a critical path of the action functional. To state the principle in more detail we first define a *variation* of a path $\gamma(t)$ with fixed endpoints.

Definition 2.5 (Variation of a critical path)**.** Let $\gamma : [a,b] \to M$ be a smooth path. A smooth *variation* of γ with endpoints fixed is a smooth map $\Gamma : [a,b] \times (-\epsilon, \epsilon) \to M$, for some $\epsilon > 0$, such that for all $t \in [a,b]$ and $s \in (-\epsilon, \epsilon)$,

$$\Gamma(t,0) = \gamma(t), \;\; \Gamma(a,s) = \gamma(a), \;\; \Gamma(b,s) = \gamma(b).$$

We denote by Γ_s and Γ_t the partial derivatives of Γ in the first and second variables, respectively. Given a Lagrangian function $L : TM \to \mathbb{R}$ and its associated action functional S, γ is a *critical path* for S if for any smooth variation Γ of γ

$$\left.\frac{d}{ds}\right|_{s=0} S[\Gamma(\cdot, s)] := \left.\frac{d}{ds}\right|_{s=0} \int_a^b L(\Gamma(t,s), \Gamma_t(t,s))\, dt = 0.$$

We call a critical path of the action functional an *L-critical path.*

Hamilton's principle of stationary action: For a path $t \mapsto \gamma(t)$ in M to represent an actual physical motion of the mechanical system $(\mathcal{B}, \mu, \Phi)$ with Lagrangian L, γ must be L-critical.

Theorem 2.1 (The Euler-Lagrange equations in local coordinates)**.** Let the L-critical path γ lie in a coordinate chart with domain U in the configuration manifold M. Let $(q_1, \dots, q_n)$ be the coordinate functions on U and $(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n)$ be the induced coordinate function on $TU \subset TM$, consisting of tangent vectors having base point in U. In these coordinates the Lagrangian has the form $L(q_1, \dots, q_n, \dot{q}_1, \dots, \dot{q}_n)$. Then γ must satisfy the n equations:

$$\frac{\partial L}{\partial q_j}(\gamma(t), \gamma'(t)) - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}(\gamma(t), \gamma'(t))\right) = 0.$$

Proof. Let $\gamma(t)$ be an L-critical path and $\Gamma(t,s)$ a variation of γ. We write, with slight (but conventional) abuse of notation

$$L(\Gamma(t,s), \Gamma_t(t,s)) = L(q_1(t,s), \dots, \dot{q}_n(t,s))$$

and observe that

$$\frac{\partial}{\partial s} L(q_1(t,s), \dots, \dot{q}_n(t,s)) = \sum_{j=1}^n \left(\frac{\partial L}{\partial q_j} - \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\right)\frac{\partial q_j}{\partial s} + \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial \dot{q}_j}\frac{\partial q_j}{\partial s}\right).$$

Setting $V_j(t) = \frac{\partial q_j}{\partial s}(t,0)$ and noting that $V_j(a) = V_j(b) = 0$, we obtain

$$0 = \left.\frac{d}{ds}\right|_{s=0} \int_a^b L(\Gamma(t,s), \Gamma_t(t,s))\, dt = \int_a^b \sum_{j=1}^n \left(\frac{\partial L}{\partial q_j} - \frac{\partial}{\partial t}\left(\frac{\partial L}{\partial \dot{q}_j}\right)\right) V_j(t)\, dt.$$

The variation vector field $(V_1(t),\dots,V_n(t))$ can be chosen arbitrarily except for the condition that it vanishes at $t=a,b$. It follows that $\frac{\partial L}{\partial q_j}-\frac{\partial}{\partial t}\left(\frac{\partial L}{\partial \dot q_j}\right)=0$ for each j and all $t\in[a,b]$. $\square$

Consider again the example of Figure 1. Let us suppose that the system is subject to gravitational potential of the form

$$V(q)=g\int_{\mathcal{B}}\Phi(q,b)\cdot e_2\,d\mu(b)$$

where g is acceleration due to gravity. This expression should be interpreted as follows. At a configuration $q\in M$ the height (relative to some arbitrarily chosen ground level) of the material point b is $\Phi(q,b)\cdot e_2$ and the amount of mass at this height around b is $d\mu(b)$. The contribution to the gravitational potential energy (near the surface of the Earth) of a particle of matter is its height times its mass times the constant g. The integral over $\mathcal{B}$ then gives the total potential energy of the body. An integral calculation using the expression 2.1 yields the physically obvious result: $V(x,[\theta])=gm_2 l\sin\theta$. (Note that θ, for the configuration shown in Figure 1, is negative and that we omit the contribution of m_1 to the potential function since the pendulum pivot remains at a constant height.) The Lagrangian function for the pendulum-on-a rail system is then

$$L(x,[\theta],\dot x,\dot\theta)=\frac{1}{2}(m_1+m_2)\dot x^2+\frac{1}{2}m_2\left[\left(\frac{r^2}{2}+l^2\right)\dot\theta^2-2\dot x\dot\theta l\sin\theta\right]-gm_2 l\sin\theta.$$

Note the term $m_2r^2\dot\theta^2/4$ in the kinetic energy. It represents the rotational energy of the bob, which rotates with the same angular speed as the rod itself.

It is now easy to obtain the equations of motion, which are the two equations (in x and θ) obtained from Theorem 2.1. They are described in the next exercise.

Exercise 2.1 (Pendulum on a rail)**.** Derive the following facts for the system of Figure 1.

1. The Euler-Lagrange equations are
   - $0=\frac{\partial L}{\partial x}-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot x}\right)=-(m_1+m_2)\ddot x+m_2 l\ddot\theta\sin\theta+m_2 l\dot\theta^2\cos\theta$
   - $0=\frac{\partial L}{\partial \theta}-\frac{d}{dt}\left(\frac{\partial L}{\partial \dot\theta}\right)=\frac{1}{2}m_2\left[\left(r^2+2l^2\right)\ddot\theta-2\ddot x l\sin\theta\right]+gm_2 l\cos\theta.$
2. Define $x_c=x+\frac{m_2}{m_1+m_2}l\cos\theta$, the horizontal component of the *center of mass* of the system. Using the first of the Euler-Lagrange equations show that $\ddot x_c=0$. Therefore, $x_c=v_c t+x_c(0)$, where v_c is the constant velocity of x_c. Therefore,

   $$x(t)=a+bt-\frac{m_2}{m_1+m_2}l\cos\theta(t)$$

   where the constants a,b are specified by a choice of initial conditions.
3. Show that θ satisfies the differential equation

   $$\left[\frac{1}{2}\left(r^2+2l^2\right)-\frac{m_2}{m_1+m_2}l^2\sin^2\theta\right]\ddot\theta-\frac{m_2}{m_1+m_2}l^2\left(\sin\theta\cos\theta\right)\dot\theta^2+gl\cos\theta=0.$$

4. Express the equation of part 3 as a system of first order equations in θ and $\dot{\theta}$. This system can be interpreted as the vector field on the plane $\mathbb{R}^2$. The corresponding field of directions (not taking into account the magnitude of the vectors) is shown in Figure 2. A few integral curves are also shown.

5. When $m_1 = \infty$ and $r = 0$, the previous equation reduces to the standard equation of motion of the simple pendulum:

$$\ddot{\theta} + \frac{g}{l}\cos\theta = 0.$$

6. If $r = 0$ and $m_1 = 0$, show that the differential equation for θ obtained in part 3 reduces to

$$\left(\ddot{y} + \frac{g}{l}\right)\cos\theta = 0$$

   where $y = \sin\theta$. Does this equation make physical sense? (If $\cos\theta \neq 0$, $\ddot{y} + g/l = 0$ implies that the bob is in free-fall. When $\cos\theta = 0$ the rigidity of rod suddenly manifests itself.)

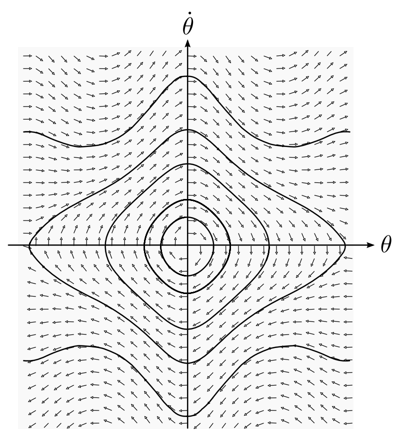

Figure 2: Expressing the second order equation in θ given in part 3 of this exercise as a system of two first order equations in the $(\theta, \dot{\theta})$-plane produces this direction field.

2.3 Example of configuration manifold with boundary

Mechanical systems on configuration manifolds with boundary, or more generally corners, are interesting in that boundary configurations represent collisions between constituent parts of the system. In this section we discuss in some detail one of the simplest examples of this kind, shown in Figure 3, which will be used to introduced some elementary but useful ideas.

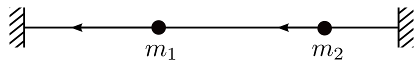

Figure 3: Two point masses can slide freely on an interval. They collide elastically with each other and with the fixed walls at the endpoints of the interval. We assume that interval has length L.

The material body in this case consists of two points, which we identify with the indices of the masses, $\mathcal{B} = \{1,2\}$, and the mass distribution measure μ is simply $\mu(i) = m_i$, $i = 1,2$. The manifold of configuration is the set

$$M = \{q = (x_1, x_2) \in [0,L] \times [0,L] : x_1 \leq x_2\}$$

which is a manifold with corners. Here L is the length of the interval. The position map is simply $\Phi(x,i) = x_i$. The kinetic energy Riemannian metric is given by

$$\langle u, v\rangle = m_1 u_1 v_1 + m_2 u_2 v_2$$

and the kinetic energy is $K(q,v) = \frac{1}{2}\|v\|^2$ where $\|\cdot\|$ is the norm associated to this Riemannian metric. Clearly, geodesics in this metric are the same as for the Euclidean metric on the plane, so the dynamic in the interior of M is described by uniform rectilinear motion.

Further assumptions about the nature of contact between the point masses and between each point mass and the walls are needed in order to determine the motion after a trajectory hits the boundary of M. A natural choice, which we make here, is to impose the conditions that the total energy and linear momentum do not change. So let q be a boundary point in M and denote by v^- and v^+ the velocities of a trajectory in M before and after, respectively, it hits the boundary at q. We first consider the case of a collision between the two point masses. First note that $\tau = (1,1)/\sqrt{m_1+m_2}$ is a unit vector (in the kinetic energy metric) tangent to the diagonal boundary component $x_1 = x_2$. The momentum associated to a velocity v at q is

$$m_1 v_1 + m_2 v_2 = m^{1/2}\langle v, \tau\rangle$$

where $m := m_1 + m_2$. Therefore, conservation of momentum in a collision between the points masses can be expressed by saying that the component of the velocity tangential to the boundary at q must remain invariant. Energy conservation, interpreted in terms of the kinetic energy metric, means that the length of the velocity vector is also invariant. This means that the velocity v^+ after collision is the mirror reflection of the velocity v^- before collision on the boundary of M. A similar but simpler argument applies to collisions of each point mass with the ends of the interval, if we imagine those ends as being walls of infinite mass and the collisions energy preserving.

Proposition 2.1 summarizes the foregoing discussion. The term *elastic collision* will be used to indicate that the assumptions made above about energy and momentum conservation at collisions are in place.

Proposition 2.1. Relative to the kinetic energy (Euclidean) Riemannian metric on the configuration space M, the motion of the two point masses system, with masses m_1 and m_2 and with elastic collisions, corresponds to billiard motion (with specular reflection at the boundary) on a right triangle shaped billiard table. The angle θ at one of the two vertices contained in the hypothenuse satisfies $\tan\theta = \sqrt{m_1/m_2}$.

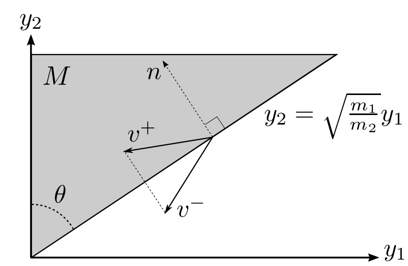

Figure 4: The configuration manifold of the two particles system described in terms of the kinetic energy Riemannian metric. Dynamics correspond to ordinary billiard motion in a right-triangle whose shape is determined by the ratio of the two masses.

An alternative way to describe the conclusion of Proposition 2.1 is as follows. Introduce a new system of coordinates in $\mathbb{R}^2$ by the expression $y_i = (m_i/m)^{1/2}x_i/l$, $i = 1, 2$. In the new system, the kinetic energy of a state $((y_1, y_2), (\dot{y}_1, \dot{y}_2))$ assumes the form $ml^2(\dot{y}_1^2 + \dot{y}_2^2)$ and energy is now proportional to Euclidean length. In this new coordinate system motion of the two-particle system is as described in the proposition. While the triangle was isosceles in the (x_1, x_2)-coordinates, in the new metric *pulled-back* under the coordinate change the sides of the triangle adjacent to the right angle have lengths $\sqrt{m_i/m}$, $i = 1, 2$. (By the choices made, lengths in the new coordinates (y_1, y_2) have been rendered *dimensionless* in the sense that they free of physical units.)

Exercise 2.2. Consider the motion of two point masses on the half-infinite line as shown in Figure 5 We suppose that the point mass m_2 starts at a point x_2 with velocity $v_2 < 0$ and that

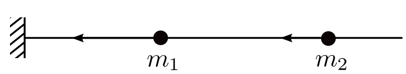

Figure 5: Two masses system in a half-infinite interval.

m_1 starts with velocity v_1. After a certain number n_c of collisions between the two masses and between m_1 and the left wall, m_2 reverses direction and moves back towards the right end of the interval, never again to hit m_1. Show that the total number n_c of collisions can be bounded from above by

$$n_c \le \left\lceil \left(\arctan \sqrt{\frac{m_1}{m_2}} \right)^{-1} \pi \right\rceil$$

Where $\lceil x \rceil$ denotes the least integer greater than or equal to x. (Suggestion: Consider the wedge shaped billiard table obtained in the way we did for the triangle of Proposition 2.1 then "unfold" it as shown in Figure 6.)

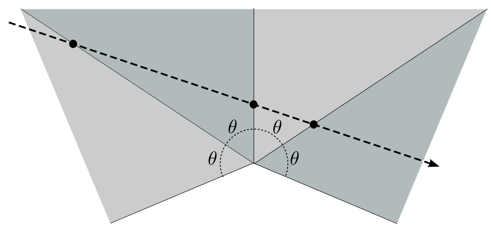

Figure 6: The *unfolding* method used in Exercise 2.2.

3 Invariant volume forms

4 Hamiltonian formalism (on TM rather than T^*M)

It turns out to be very useful to study the equations of motion of a mechanical system on its phase space, or position-velocity space, TM. This is mainly because Newton's second order differential equation on M becomes a first order equation on TM; but also because the rich geometric structure of the tangent bundle can be used to derive general properties of the system in a particularly elegant and often straightforward way. This is even more so for the cotangent bundle, or position-momentum space T^*M, although we will focus on TM to a greater extent in these notes.

Recall that the Euler-Lagrange equations express the equations of motion of a mechanical system on TM in local coordinates $(q_1,\ldots,q_n,\dot{q}_1,\ldots,\dot{q}_n)$ on the tangent bundle. Our immediate goal is to find a coordinate-free expression of the same equations. The advantages of doing so will, hopefully, become apparent soon. But first, we need to introduce several definitions related to the geometric structure of TM.

4.1 The Lagrangian one-form and energy function

We write $N := TM$ for simplicity. The base point projection onto M will be written $\pi : N \to M$ and $N_q := \pi^{-1}(q) = T_qM$ will denote the vector space fiber above $q \in M$. We often denote elements of N by v, instead of the more explicit (q,v) used earlier, and write $q = \pi(v)$. Being itself a manifold, N has its own tangent bundle $TN = T(TM)$. The kernel V_v of the differential $d\pi_v : T_vN \to T_qM$ is a vector space of dimension $n = \dim M$, which is canonically isomorphic to T_qM via the linear isomorphism $\mathcal{I}_v : T_qM \to V_v$ constructed as follows. For each $w \in T_qM$ consider the path $\xi(t) = v + tw$ in N_q. Then $\mathcal{I}_v(w) := \xi'(0)$. It is clear that $\mathcal{I}_v(w)$ projects to 0 under $d\pi_v$, so it belongs to V_v, and that it is injective, hence an isomorphism since V_v and T_qM have the same dimension. Let now L be a real valued smooth function on N. The *fiber derivative* of L at $v \in N$ is defined by $\mathbb{F}L_v(w) := dL_v(\mathcal{I}_v(w))$. It will be useful to note that $v \mapsto \mathbb{F}L_v$ is a map from TM to T^*M.

Definition 4.1 (Canonical vector field, Lagrangian 1-form, and energy function)**.** The vector field on N defined by $v \mapsto \mathcal{Z}_v := \mathcal{I}_v(v)$ will be referred to as the *canonical vector field* of TM. Let $L : N \to \mathbb{R}$ be any smooth function, which we call a *Lagrangian* function on N. The Lagrangian one-form $\Theta = \Theta_L$ associated to L is defined at $v \in N$ by $\Theta_v := dL_v \circ \mathcal{I}_v \circ d\pi_v$. The *energy function* $E = E_L : N \to \mathbb{R}$ associated to L is defined by $E(v) = \mathcal{Z}L - L$, where $\mathcal{Z}L = dL(\mathcal{Z})$ is the derivative of L in the direction of the canonical vector field.

Let us illustrate these definitions for a Newtonian system on a Riemannian manifold $(M, \langle\cdot,\cdot\rangle)$.

Proposition 4.1 (Lagrangian one-form and energy function for Newtonian systems)**.** Let M be a Riemannian manifold with metric $\langle\cdot,\cdot\rangle$ and Lagrangian $L(v) = \frac{1}{2}\|v\|^2 - U(\pi(v))$. Then the Lagrangian one-form has the form

$$\Theta_v(\xi) = \langle v, d\pi_v \xi\rangle$$

for all $v \in N = TM$ and $\xi \in T_vN$. The energy function for the same system is the sum of the potential and kinetic energies:

$$E(v) = \frac{1}{2}\|v\|^2 + U(\pi(q)).$$

The fiber derivative of L at $v \in N$ is $\mathbb{F}L_v(w) = \langle v, w\rangle$.

Proof. * $\square$

Exercise 4.1 (Coordinate expressions of E and Θ_L)**.** Let $(q_1, \dots, q_n, \dot q_1, \dots, \dot q_n)$ be a coordinate system on $N = TM$ and $L(q_1, \dots, q_n, \dot q_1, \dots, \dot q_n)$ a smooth function on N. Show that the following hold.

1. The Lagrange one-form Θ is given in coordinates by $\Theta_L = \sum_j \frac{\partial L}{\partial \dot q_j} dq_j$.

2. The energy function is $E = \sum_j \dot q_j \frac{\partial L}{\partial \dot q_j} - L$.

3. Let $\gamma(t)$ be a smooth path in M, so that $\gamma'(t)$ is a smooth path in N. Then $\gamma''(t)$ makes sense as a vector field along the path $t \mapsto \gamma'(t) \in N$. Let $d\Theta$ be the exterior derivative of Θ. Show that
$$\gamma'' \lrcorner\, d\Theta + dE = \sum_j \left(\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q_j}\right) - \frac{\partial L}{\partial q_j}\right) dq_j$$
where $(u \lrcorner\, d\Theta)(v) := d\Theta(u, v)$.

4. Let $u_j(t) := q_j(\gamma_j(t))$. Show that

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot q_j}\right) - \frac{\partial L}{\partial q_j} = \sum_k \left(\frac{\partial^2 L}{\partial \dot q_j \partial q_k} u_k' + \frac{\partial^2 L}{\partial \dot q_j \partial \dot q_k} u_k''\right) - \frac{\partial L}{\partial q_j}.$$

Proposition 4.2 (Coordinate-free expression of the Euler-Lagrange equations)**.** A smooth path $\gamma : [a,b] \to M$ is L-critical with fixed endpoints if and only if the path $\eta(t) := \gamma'(t)$ in N is a solution of the differential equation $\eta' \lrcorner\, d\Theta + dE = 0$.

Proof. This is an immediate consequence of Exercise 3.1. □

Corollary 4.1 (Conservation of energy)**.** If γ is an L-critical path with fixed endpoints, $E(\gamma'(t))$ is constant in t.

Proof. Taking the derivative in t

$$\frac{d}{dt}E(\gamma'(t)) = dE_{\gamma'(t)}(\gamma''(t)) = -(\gamma''(t) \lrcorner\, d\Theta)(\gamma''(t)) = -d\Theta(\gamma''(t), \gamma''(t)) = 0.$$

Therefore, $E(\gamma'(t))$ is constant as claimed. □

We have seen in Proposition 3.2 that for a path $\gamma(t)$ on M to be L-critical it is necessary that γ satisfy the second order differential equation $\gamma'' \lrcorner\, d\Theta = -dE$. For this to be a regular second-order differential equation, one must be able locally to solve for the second derivative γ''. Looking at part (4) of Exercise 3.1 we see that in order to solve for γ'' the matrix $(L_{\dot q_i \dot q_j})$ of second derivatives of L in the velocity coordinates $\dot q_j$ must be invertible. If this is the case we say that L is *regular*. For example, if the Lagrangian has the form

$$L(q,v) = \frac{1}{2}Q_q(v) - U(q)$$

where $Q_q(v)$ is a quadratic function in v whose coefficients are functions of q, then by a simple computation we obtain, for a vector $w = (w_1, \dots, w_n)$ in T_qM,

$$\sum_{ij} L_{\dot q_i \dot q_j}(q,v) w_i w_j = Q_q(w).$$

Thus L is regular in this case exactly when Q is a non-degenerate quadratic form. This is always the case for a Newtonian Lagrangian, for which the kinetic energy term is the Riemannian metric of the configuration manifold M.

Exercise 4.2 (Regular Lagrangian)**.** Let L be a smooth real valued function on the n-dimensional manifold M. Show that the following conditions are equivalent.

1. In every coordinate system $(q_1, \dots, q_n, \dot q_1, \dots, \dot q_n)$ of TM the n-by-n matrix $(L_{\dot q_i \dot q_j}(q, \dot q))$ is invertible for all $(q, \dot q)$.
2. In every coordinate system $(q_1, \dots, q_n, \dot q_1, \dots, \dot q_n)$ of TM, the functions on TM

$$q_1, \dots, q_n, \frac{\partial L}{\partial \dot q_1}, \dots, \frac{\partial L}{\partial \dot q_n}$$

have everywhere linearly independent differentials.

3. In every coordinate system $(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n)$ of TM, defining

$$f_j := q_j, \;\; g_j := \frac{\partial L}{\partial \dot{q}_n},$$

then $(f_1, \ldots, f_n, g_1, \ldots, g_n)$ is also a local coordinate system of TM.

4. The two-form $\omega_L := d\Theta_L$, where Θ_L is the Lagrangian one-form for L, is non-degenerate.

5. The map $\mathbb{F}L : TM \to T^*M$ is locally invertible.

If L satisfies any of these equivalent conditions, we say that L is *regular*. If $\mathbb{F}L$ is a diffeomorphism from TM to T^*M, then L is said to be *hyperregular*.

Exercise 4.3. We assume here the same notations used in Exercise 3.2, in particular the functions f_j, g_j and the two-form $\omega := \omega_L$. Show that

1. $\Theta = \sum_j g_j dq_j$

2. $\omega = -\sum_j df_j \wedge dg_j$

3. $\xi \lrcorner \omega = \sum_j \left(dg_j(\xi) df_j - df_j(\xi) dg_j \right)$ for any $\xi \in TN$

4. $\mathbb{F}L_v \frac{\partial}{\partial \dot{q}_j} = g_j(v)$

5. Define on T^*M the canonical one-form α such that $\alpha_\eta(X) = \eta(d\pi_\eta X)$, where $\eta \in T^*M$ and π is the base-point map of $T^*M \to M$. Show that $\mathbb{F}L^* \alpha = \Theta$.

The functions g_j, when the Lagrangian is regular, are called the *momentum* coordinates, although more often one reserves the term for the functions $p_j = g_j \circ (\mathbb{F}L)^{-1}$ defined locally on T^*M. In this case $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ constitute a system of local coordinates of T^*M. We nevertheless use $p_j := g_j$ in places below and call these functions the *conjugate momentum* coordinates to the q_j.

4.2 The Hamiltonian vector field

Proposition 4.3 (The Hamiltonian vector field)**.** Let Θ and E be the Lagrange one-form and energy function associated to a regular Lagrangian L. In particular, the closed two-form $\omega = d\Theta$ is non-degenerate. Let X be the unique vector field on TM defined by the equation $X \lrcorner \omega = -dE$ and $\pi : TM \to M$ the base point projection. Then

1. $d\pi_v X = v$ for all $v \in TM$;

2. If a smooth path γ on M is L-critical then the path γ' in TM is an integral curve of X;

3. If η is an integral curve of X then $\gamma = \pi \circ \eta$ is L-critical.

Proof. * □

Exercise 4.4 (Hamiltonian vector field in coordinates)**.** Let a given regular Lagrangian be expressed in local coordinates by $L(q_1,\dots,q_n,\dot q_1,\dots,\dot q_j)$ and let $p_j := \frac{\partial L}{\partial \dot q_j}$ be the conjugate momentum coordinates to the q_j. We know that for a regular Lagrangian $(q_1,\dots,q_n,p_1,\dots,p_n)$ also defines a system of local coordinates on $N = TM$. In these new coordinates, write the energy function as $H(q_1,\dots,q_n,p_1,\dots,p_n) = E(q_1,\dots,q_n,\dot q_1,\dots,\dot q_j)$. Show the following.

1. $\omega = \sum_j dp_j \wedge dq_j$;
2. $X = \sum_j \left(\frac{\partial H}{\partial p_j}\frac{\partial}{\partial q_j} - \frac{\partial H}{\partial q_j}\frac{\partial}{\partial p_j}\right)$;
3. An L-critical path $\gamma(t) = (q_1(t),\dots,q_n(t),p_1(t),\dots,p_j(t))$ satisfies the system of *Hamilton's equations*
$$\dot q_j = \frac{\partial H}{\partial p_j}, \quad \dot p_j = -\frac{\partial H}{\partial q_j}$$
for $j = 1,\dots,n$.

4.3 Lagrangian symmetries and conservation laws

We restrict attention to mechanical systems of the Newtonian type; that is, systems whose configuration manifold is a Riemannian manifold having a Lagrangian function of the form $L(q,v) = \frac{1}{2}\|v\|^2 - U(q)$, where U is the potential function.

Definition 4.2 (Symmetries of the Lagrangian)**.** A *symmetry* of the Newtonian Lagrangian is a smooth diffeomorphism of M that leaves L invariant. More precisely, a diffeomorphism $\varphi : M \to M$ is a symmetry of L if

$$L(\varphi(q), d\varphi_q v) = L(q,v)$$

for all (q,v) in TM. We say that a complete smooth vector field Y on M is an *infinitesimal symmetry* of L if its flow Φ_t is a one-parameter group of symmetries of L.

Exercise 4.5. Show that a diffeomorphism φ of the Riemannian manifold M is a symmetry of the Newtonian Lagrangian $L(q,v) = \frac{1}{2}\|v\|^2 - U(q)$ if and only if φ is an isometry of M that leaves invariant the potential energy function U.

Proposition 4.4. Let Y be a complete smooth vector field on a Riemannian manifold M with associated flow Φ_s, and L a Newtonian Lagrangian function on $N = TM$ with energy function E and Lagrangian one-form Θ. For any given smooth path $\gamma : [a,b] \to M$, we define $S(s) := \int_a^b L((\Phi_s \circ \gamma)'(t))\, dt$. Then

$$S'(0) = \left\langle \gamma'(b), Y_{\gamma(b)} \right\rangle - \left\langle \gamma'(a), Y_{\gamma(b)} \right\rangle - \int_a^b \left\langle \frac{\nabla \gamma'}{dt} + \operatorname{grad} U, Y_{\gamma(t)} \right\rangle dt.$$

In particular, if $\gamma(t)$ is an L-critical path and Y is an infinitesimal symmetry of L, then $S'(0) = 0$, the integrand vanishes, and $\left\langle \gamma'(b), Y_{\gamma(b)} \right\rangle = \left\langle \gamma'(a), Y_{\gamma(b)} \right\rangle$.

Proof. * □

Corollary 4.2 (Noether's theorem on symmetries and conservation laws)**.** Let a smooth vector Y field on a Riemannian manifold M be an infinitesimal symmetry of the Newtonian Lagrangian $L : TM \to \mathbb{R}$. Then the quantity $\mathcal{P}_Y(q,v) := \langle v, Y_q\rangle$ is a constant of motion. In other words, if an L-critical path $\gamma(t)$ in M represents the motion of a mechanical system with Lagrangian L, then $t \mapsto \mathcal{P}_Y(\gamma(t),\gamma'(t))$ is constant. We call the function $\mathcal{P}_Y : TM \to \mathbb{R}$ the *momentum* associated to the infinitesimal symmetry Y.

Proof. This is a corollary of Proposition 3.4. □

Let us consider one standard example. A Newtonian system of n point masses in $\mathbb{R}^3$ has the Lagrangian function

$$L(x,v) = \sum_{j=1}^{n} \frac{1}{2} m_j |v_j|^2 - U(x_1,\dots,x_n)$$

where the potential energy is of the form

$$U(x) = \sum_{i<j} V_{ij}(|x_j - x_i|).$$

The Euclidean group $SE(3)$ acts on the configuration manifold $M = \underbrace{\mathbb{R}^3 \times \cdots \times \mathbb{R}^3}_{n}$ by

$$g(x_1,\dots,x_n) = (gx_1,\dots,gx_n).$$

Since each g leaves invariant the norm of vectors and the distance between points, it acts by symmetries of the Lagrangian. Now, let $Z = (X,w) \in \mathfrak{se}(3)$ be an element of the Lie algebra of the Euclidean group, where $X \in \mathfrak{so}(3)$ and $w \in \mathbb{R}^3$. We also denote by Z the vector field induced on $\mathbb{R}^3$ by Z, which has the form $Z_x = Xx + w$, where Xx is interpreted as the matrix multiplication of the skew-symmetric matrix X and the column vector $x \in \mathbb{R}^3$, and w is a column vector in $\mathbb{R}^3$. The vector field induced by Z on M is then

$$Z_x = (Xx_1 + w,\dots,Xx_n + w).$$

The momentum function associated to $Z = (X,w) \in \mathfrak{se}(3)$, evaluated at a state (x,v) is then

$$\mathcal{P}_Z(x,v) = \sum_{j=1}^{n} m_j v_j \cdot (Xx_j + w), \tag{4.1}$$

where we recall that the Riemannian metric on M is given by $\langle v,u\rangle = \frac{1}{2}\sum_{j=1}^{n} m_j v_j u_j$.

The expression 3.1 can be made to look more familiar as follows. We have seen that to each $X \in \mathfrak{so}(3)$ is associated an $a \in \mathbb{R}^3$ such that $Xx = a \times x$. Furthermore, the familiar triple product $v \cdot (a \times x)$ in vector calculus satisfies $v \cdot (a \times x) = -a \cdot (v \times x)$. Therefore, equality 3.1 can be stated as follows. For all $a_j, w \in \mathbb{R}^3$, the quantity

$$\mathcal{P}_Z(x,v) = w \cdot \left(\sum_{j=1}^{n} m_j v_j\right) - a \cdot \left(\sum_{j=1}^{n} m_j v_j \times x_j\right)$$

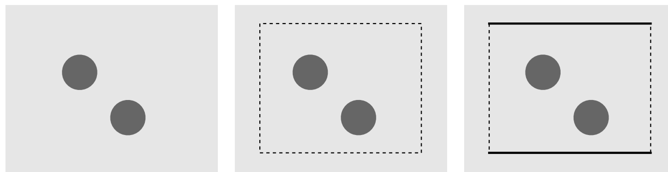

Figure 7: A system of two free discs interacting via elastic collisions. On the left: the ambient space is $\mathbb{R}^2$ and the symmetry group is $SE(2)$. In the middle: periodic boundary conditions are imposed and the two discs now lie in a torus. In this case the symmetry group reduces to $\mathbb{R}^2$ acting by translations. On the right, two reflecting, perfectly smooth walls (so that no momentum is exchanged between the walls and the discs tangentially to the walls) are introduced. The motion now is in a cylinder with reflecting circle boundaries. In this case the symmetry group reduces to $\mathbb{R}$, acting by translations in the horizontal direction. In the first case, linear momentum in the x and y directions and total angular momentum are conserved. In the middle case only the two components of linear momentum are conserved; and in the last case only the x-component of linear momentum is conserved.

is constant along trajectories of the system. Therefore,

$$P := \sum_{j=1}^{n} m_j v_j, \quad L := \sum_{j=1}^{n} m_j x_j \times v_j$$

are, separately, constant. The quantity P is called the *linear momentum*, and L the *angular momentum* of the particle system.

In other cases, the Lagrangian may be invariant under a subgroup of $SE(3)$, rather than the full group. For example, suppose that a pendulum in $\mathbb{R}^3$ with a rigid rod and pivot fixed at the origin, is subject to a constant acceleration due to gravity. The pivot being fixed reduces the symmetry group to a subgroup of $SO(3)$, and gravity reduces it further to the group $SO(2)$ of rotations about the axis in the direction parallel to the force of gravity through the pivot. Therefore, only the single quantity $e \cdot (v \times x)$ is conserved, where x is the position of the bob, v is its velocity, and e is the direction of gravity.

4.4 The symplectic form and Sasaki metric on TM

Let M be a Riemannian manifold with metric $\langle \cdot, \cdot \rangle$. Let ∇ be the Levi-Civita connection and $\mathcal{K}$ the connection map on the tangent bundle of $N = TM$. Recall that $\mathcal{K}$ and ∇ are related as follows. For $v \in N$ and $\xi \in T_v N$, let $v(t)$ be a path in N representing ξ; that is, $v(0) = v$ and $v'(0) = \xi$. Then

$$\mathcal{K}_v \xi := \left. \frac{\nabla}{dt} \right|_{t=0} v(t).$$

We can similarly express the connection map in terms of the covariant derivative of vector fields as follows. Let Y be a smooth vector field defined on a neighborhood of a point q in M such that $Y(q) = v$. Let w be any vector in $T_q M$. Then $\xi := dY_q w$ is a vector in $T_v N$ projecting to w under the differential of the base-point map $\pi : N \to M$ (due to the chain rule and the fact that $\pi \circ Y$ is the identity map on M) and

$$\mathcal{K}_v \xi = \nabla_w Y.$$

Also recall that TN splits as a direct sum of subbundles $TN = H \otimes V$ where, at each $v \in N$, the horizontal subspace H_v is the kernel of $\mathcal{K}_v : T_v N \to T_q M$, $q = \pi(v)$, and V_v is the vertical subspace, defined as the kernel of $d\pi_v : T_v N \to T_q M$. The projection $d\pi_v$ restricted to H_v is a linear isomorphism onto $T_q M$ and recall that the map $\mathcal{I}_v : T_q M \to V_v$ is a linear isomorphism such that $\mathcal{K}_v \circ \mathcal{I}_v w = w$ for all $w \in T_q M$.

Proposition 4.5. Let L be the Lagrangian function on N associated to a Newtonian system with a given potential function U on the Riemannian manifold M. Therefore, $L(q, v) = \frac{1}{2}\|v\|^2 - U(q)$. Let $\Theta = \Theta_L$ be the associated Lagrangian one-form. Then

$$d\Theta(\xi_1, \xi_2) = \langle \mathcal{K}_v \xi_1, d\pi_v \xi_2 \rangle - \langle \mathcal{K}_v \xi_2, d\pi_v \xi_1 \rangle$$

for all $v \in N$ and $\xi_1, \xi_2 \in T_v N$.

Proof. * □

Definition 4.3 (Sasaki metric)**.** The *Sasaki metric* on N is the Riemannian metric on N defined on tangent vectors $\xi_1, \xi_2 \in T_v N$ by

$$\langle \xi_1, \xi_2 \rangle_v := \langle d\pi_v \xi_1, d\pi_v \xi_2 \rangle_q + \langle \mathcal{K}_v \xi_1, \mathcal{K}_v \xi_2 \rangle_q.$$

In other words, the Sasaki metric is characterized by the requirements that H and V be orthogonal subbundles of TN, that the projection $d\pi_v$ be an isometric isomorphism between H_v and $T_q M$, and that $\mathcal{K}_v$ be an isometric isomorphism between V_v and $T_q M$.

It is interesting to note that the symplectic form $\omega = d\Theta$ and the Sasaki metric $\langle \cdot, \cdot \rangle$ are related by $\omega(\cdot, \cdot) = \langle J \cdot, \cdot \rangle$, where $v \mapsto J_v$ is a field of linear isomorphisms on $T_v N$ such that $J_v^2 = -I$ for each v. This means that N has a natural quasi-complex structure. When ∇ is flat, it can be shown that the quasi-complex structure is integrable and N is a complex manifold.

Definition 4.4 (Vertical and horizontal lifts of vector fields)**.** Let Y be a vector field on M. We denote by Y^H the unique vector field on N such that $Y_v^H \in H_v$ and $d\pi_v Y_v^H = Y_q$, $q = \pi(v)$; and by Y^V the unique vector field on N such that $Y_v^V = \mathcal{I}_v Y_q$.

Definition 4.5 (Gradient of a function)**.** Let f be a smooth function on a Riemannian manifold M with Riemannian metric $\langle \cdot, \cdot \rangle$. The *gradient vector field* of f is the vector field on M uniquely determined by the equation $\langle \text{grad}_q f, u \rangle = df_q(u.)$

Proposition 4.6. Let X be the Hamiltonian vector field in N associated to a Newtonian Lagrangian function L with potential energy function U. Let E be the energy function derived from L, X_0 the Hamiltonian vector field on the same space associated to free motion (that is, $U = 0$), which is called the *geodesic spray* of the Riemannian metric on M. Finally, let $\mathcal{Z}$ be canonical vector field on N. The gradient vector field of a function on N is taken with respect to the Sazaki metric. Then $X = X_0 - (\text{grad } U)^V$, $\text{grad}\, E = \mathcal{Z} + (\text{grad}\, U)^H$ and $\text{grad}\, L = \mathcal{Z} - (\text{grad}\, U)^H$.

Proof. First note that $X - X_0$ is a vertical vector field since both X and X_0 project under $d\pi_v$ to v. Since X_0 is horizontal, $(X - X_0)_v = \mathcal{K}_v X_v$. But this last expression is the acceleration vector , at time $t = 0$, of a solution path to Newton's equation with initial condition (q, v). Therefore, it is equal to $-\text{grad}_q U$. This shows the first identity. For the second identity first observe that

$$dE_v(\xi) = \omega_v(\xi, X) = \langle \mathcal{K}_v \xi, d\pi_v X\rangle - \langle \mathcal{K}_v X, d\pi_v \xi\rangle = \langle \xi, \mathcal{Z}_v\rangle + \langle \xi, (\text{grad}\, U)^H\rangle.$$

The second identity follows from this. An easy calculation shows that the gradient of the kinetic energy function $K(q, v) := \frac{1}{2}\|v\|_q^2$ is $\mathcal{Z}$. The third identity is then a consequence of $E + L = 2K$. $\square$

A consequence of the proposition is that

$$\|\text{grad}_v E\|_v^2 = \|\text{grad}_v L\|_v^2 = \|v\|_q^2 + \|\text{grad}_q U\|_1^2.$$

As the energy function E is invariant under the Hamiltonian flow, level sets of E are invariant sets. When $\mathcal{E}$ is a regular value of E, the equation $E = \mathcal{E}$ defines a smooth manifold, possibly with boundary, of dimension $2n - 1$, where n is the dimension of M. This is a consequence of the implicit function theorem. The vector field

$$\eta := \text{grad}\, E/\|\text{grad}\, E\|$$

is a unit normal vector field to the level sets $N(\mathcal{E})$. The tangent bundle to $N(\mathcal{E})$ is the kernel of the one-form dE; equivalently, the orthogonal space to η. It follows that

$$T_v N(\mathcal{E}) = \{\xi \in T_v N : d(U \circ \pi)_v \xi + \langle v, \mathcal{K}_v \xi\rangle = 0\}.$$

The restriction of π to $N(\mathcal{E})$ is a *submersion*, that is, $d\pi_v$ maps each tangent space $T_v N(\mathcal{E})$ onto $T_q M$, $q = \pi(v)$. In fact, let $w \in T_q M$ and denote by $\overline{w}$ the horizontal lift of w to v and define

$$\xi = \overline{w} - \frac{dU_q(w)}{\|v\|^2} Z_v.$$

Then $d\pi_v \xi = w$ and

$$d(U \circ \pi)_v \xi + \langle v, \mathcal{K}_v \xi\rangle_q = dU_q(w) + \left\langle v, -\|v\|_q^{-2} dU_q(w) v\right\rangle_q = 0$$

Let now S be a smooth hypersurface, or codimension one submanifold, of M. We denote $N_S = \{(q, v) \in N : q \in S\}$ and define $N_S(\mathcal{E}) := N_S \cap N(\mathcal{E})$. Note that for q to be the base point of a state in $N_S(\mathcal{E})$ it is necessary that $\mathcal{E} \geq U(q)$.

Proposition 4.7. Let $\mathcal{E}$ be a regular value of E and S a smooth hypersurface of M, which may be an open subset of regular points of the boundary of M. Then $\xi \in T_v N_S(\mathcal{E})$ if and only if

$$d(U \circ \pi)_v \xi + \langle v, \mathcal{K}_v \xi\rangle = 0$$

and $d\pi_v \xi$ is tangent to S at $\pi(v)$. In particular, the subspaces

$$\{\xi \in V_v : \langle \mathcal{Z}, \xi\rangle_v = 0\} \text{ and } \{\xi \in H_v \cap \ker d(U \circ \pi)_v : d\pi_v \xi \in T_q S\}$$

lie in $T_v N_S(\mathcal{E})$.

Proof. * □

Proposition 4.8. Let $\omega = d\Theta$ be the symplectic form on $N = TM$. Let i be the inclusion map of $N_S(\mathcal{E})$ into N and define $\omega_S^{\mathcal{E}} := i^*\omega$. Then $\omega_S^{\mathcal{E}}$ is a symplectic form on $N_S(\mathcal{E}) \setminus TS$.

Proof. The form $\omega_S^{\mathcal{E}}$ is the restriction of ω to vectors tangent to $N_S(\mathcal{E})$. There will be little risk of confusion in denoting it simply by ω. This form is closed since d commutes with i^*, so we only need to show that it is also non-degenerate. Let ξ be a vector in $T_v N_S(\mathcal{E})$ and assume that $\omega(\xi,\zeta) = 0$ for all $\zeta \in T_v N_S(\mathcal{E})$ and $v \notin TS$. We first assume that ζ is vertical, so $\langle v, \mathcal{K}_v\zeta\rangle = 0$ by Proposition 3.7 and $\omega(\xi,\zeta) = -\langle d\pi_v\xi, \mathcal{K}_v\zeta\rangle = 0$, where we use the form of ω given by Proposition 3.5. This means that $\langle w, d\pi_v\xi\rangle = 0$ for all $w \perp v$, hence $d\pi_v\xi = \lambda v$ for some real number λ. Since also $d\pi_v\xi \in T_qS$ by Proposition 3.7 and $v \notin T_qS$, we conclude that $d\pi_v\xi = 0$ and ξ is a vertical vector. Therefore, $0 = \omega(\xi,\zeta) = \langle \mathcal{K}_v\xi, d\pi_v\zeta\rangle$ for all $\eta \in T_v N_S(\mathcal{E})$. Because the restriction of $d\pi_v$ to $T_v N_S(\mathcal{E})$ maps onto T_qS as noted in the comments preceding Proposition 3.7, we have $\langle \mathcal{K}_v\xi, w\rangle = 0$ for all $w \in T_qS$. So we have $\mathcal{K}_v\xi = \lambda n_q$, $\lambda \in \mathbb{R}$, where n_q is a unit vector perpendicular to S at q, and $d\pi_v\xi = 0$. Once again by Proposition 3.7, $0 = \langle v, \mathcal{K}_v\xi\rangle = \lambda\langle v, n_q\rangle$. Therefore, $\lambda = 0$ as $\langle v, n_q\rangle \neq 0$, and $\xi = 0$. □

Corollary 4.3. Let S be a smooth hypersurface in the n-dimensional configuration manifold M. Let $N_S(\mathcal{E})$, as defined above, be the set of (q,v) in an energy level set $N(\mathcal{E})$, for a regular value $\mathcal{E}$, such that $q \in S$. Then $N_S(\mathcal{E})$ is a symplectic manifold of dimension $2n-2$ with symplectic form given by restriction to $N_S(\mathcal{E})$ of $d\Theta$. The hypersurface may be a smooth open subset of the boundary of M.

The next exercise shows the various concepts introduced in this section in the special case in which M is submanifold of dimension n of $\mathbb{R}^n$ with smooth boundary and Riemannian metric given by the standard inner product in $\mathbb{R}^n$, denoted by $\langle u, v\rangle = u \cdot v$. Then $N := TM \cong M \times \mathbb{R}^n$ and each element of TN may be written as $(q, v, \dot{q}, \dot{v})$, which decomposes as a sum of a horizontal component $(q, v, \dot{q}, 0) \in H_{(q,v)}$ and a vertical component $(q, v, 0, \dot{v}) \in V_{(q,v)}$.

Exercise 4.6. Let $M = \mathbb{R}^n$ with Lagrangian $L(q,v) = \frac{1}{2}|v|^2$, where $|v|^2 = v \cdot v$. Show the following.

1. The Hamiltonian vector field is $X = (q, v, v, 0)$.
2. The Lagrangian one-form is $\Theta = \sum_{j=1}^n v_j\, dq_j$.
3. The symplectic form is $\omega = \sum_{j=1}^n dv_j \wedge dq_j$.
4. The Sasaki metric on $T\mathbb{R}^n$ is $\sum_{j=1}^n \left(dq_j \otimes dq_j + dv_j \otimes dv_j\right)$
5. The gradient vector field of the energy function is $\mathrm{grad}_{(q,v)} E = (q, v, 0, v)$
6. The unit normal vector field to the level sets of E is $\eta_{(q,v)} := \frac{\mathrm{grad}_{(q,v)} E}{|\mathrm{grad}_{(q,v)} E|} = \left(q, v, 0, \frac{v}{|v|}\right)$.

5 Invariant measures

An essential fact concerning the dynamics of conservative mechanical systems is the existence of volume measures on the phase space, and on level sets of the energy function, that are invariant under the Hamiltonian flow. When the volume of an energy level set is finite one obtains by normalization a natural notion of invariant probability measure. In this section we wish to describe such invariant measures as well as associated invariant measures on the boundary of the phase space energy level sets. In the next section, on ergodic theory, we will derive a number of dynamical conclusions from the existence of these measure.

We consider Newtonian systems on general Riemannian manifolds. Thus let M be an n-dimensional Riemannian manifold, which we allow to have non-empty boundary. The boundary will be denoted S. We assume that the S is a smooth manifold of dimension $n-1$, although later we may relax this smoothness assumption and consider piecewise smooth boundaries more generally. The Lagrangian is then $L(q,v) = \frac{1}{2}\|v\|^2 - U(q)$ for a smooth potential function U. In systems of billiard type we let $U = 0$. The phase space is $N = TM$, also a manifold with boundary. The energy function and Lagrangian 1-form are denoted, respectively, E and Θ, and the symplectic form ω. The Hamiltonian vector field will be denoted X and the the Hamiltonian flow, which is well-defined in the interior of N, will be denoted Φ_t.

It is necessary to specify how trajectories are to be extended past a time of collision with the boundary. Equivalently, we need to specify how a solution $\gamma(t)$ to Newton's equation $\frac{\nabla\gamma'}{dt} = -\text{grad}\, U$ (a path in M) should be defined after it reaches a boundary point. There are different, physically meaningful, choices for boundary condition, but in these notes we will only consider the following one.

Assumption 1 (Boundary condition)**.** When a solution path of Newton's equation reaches a point q in the boundary S of M, its velocity v^+ immediately after collision will be taken to be the specular reflection of the velocity v^- just prior to collision on the tangent plane to the boundary at q.

Let us take a moment to reflect on the physical meaning of this assumption. Consider two convex rigid bodies, $\mathcal{B}_1$ and $\mathcal{B}_2$ in in $\mathbb{R}^3$ moving freely in space, and colliding with each other at a given moment. The configuration manifold consists of a subset M of $G_1 \times G_2$, where $G_i = SE(3)$ is the group of Euclidean motions of body $\mathcal{B}_i$. If the bodies have smooth boundary and are convex, M is a manifold of dimension 12 having a boundary S, which is a smooth manifold of dimension 11. A boundary point of M corresponds to a collision configuration. A tangent vector at a boundary point describes a pre-collision velocity if it points outward (away from M), or a post-collision velocity if it points inward. Vectors T_qS represent grazing motion. If we assume that the boundary surfaces of the the bodies are perfectly polished and physically smooth (rather than, say, rubbery) and that the bodies do not offer any resistance to motion in this tangential direction, then the post-collision velocities of the bodies, represented by the vectors $v^\pm \in T_qM$, are determined as follows. If the map $R : v^- \mapsto v^+$ is linear, which we assume to be the case, then by conservation of energy R must be an isometric isomorphism of T_qM. (Recall that the Riemannian metric on M is derived from the kinetic energy.) In particular, $\|v^-\| = \|v^+\|$. The decomposition of a vector $v \in T_qM$ as a sum of a component in T_qS and a component in the direction perpendicular to T_qS corresponds to the decomposition of the state of the system at the moment of collision into a grazing motion and a frontal collision motion. The tangential components of the vectors v^+ and v^- are the same, due to the assumption of physical smoothness. Their components in the direction perpendicular to the boundary must switch signs because v^- points out of M and v^+ points into M. Therefore, R must be a reflection map. Note that other choices for the collision map at boundary points compatible with conservation of linear momentum, angular momentum, energy, and time reversibility are possible. We nevertheless restrict attention here to specular reflections only.

With Assumption 1, we have a well-defined way of extending trajectories for times $t > t_c$ for any collision time t_c. In this way, at least in cases when the potential function U is zero (but in many other cases as well), we have that trajectories are defined for all time. (We will not be concerned with collisions at corners and other singular events, which will typically have probability 0.)

5.1 The Liouville measure

We begin by summarizing a few general facts and the notation we will use in this section. We restrict attention to Newtonian systems, for which the Lagrangian function is given by kinetic energy minus potential energy.

- M is the configuration manifold with the kinetic energy Riemannain metric $\langle\cdot,\cdot\rangle$. The base-point projection map is $\pi : N := TM \to M$. The Lagrangian function is

$$L(q,v) = \frac{1}{2}\|v\|_q^2 - U(q)$$

  where U is a potential function on M. The associated energy function is given by $E(q,v) = \frac{1}{2}\|v\|_q^2 + U(q)$. The Lagrangian one-form is Θ which, for Newtonian systems, is given by

$$\Theta(\xi) = \langle v, d\pi_v \xi\rangle_q.$$

  The symplectic form on N is denoted by $\omega := d\Theta$.

- For each $(q,v) \in N$, the vertical lift is the previously defined isomorphism $\mathcal{J}_v$ from T_qM to the vertical subspace $V_v \subset T_vN$. The connection map $\mathcal{K}_v$ goes from T_vN to T_qM, $\mathcal{K}_v \circ \mathcal{J}_v$ is the identity map on T_qM, and the kernel of $\mathcal{K}_v$ is the horizontal subspace $H_v \subset T_vN$. The symplectic form can be expressed in terms of the connection map and the Riemannian inner product as

$$\omega(\xi_1,\xi_2) = \langle \mathcal{K}_v\xi_1, d\pi_v\xi_2\rangle - \langle \mathcal{K}_v\xi_2, d\pi_v\xi_1\rangle$$

  for all $v \in N$ and $\xi_1, \xi_2 \in T_vN$. The Sasaki metric is the Riemannian metric on N defined on tangent vectors $\xi_1, \xi_2 \in T_vN$ by

$$\langle \xi_1,\xi_2\rangle_v := \langle d\pi_v\xi_1, d\pi_v\xi_2\rangle_q + \langle \mathcal{K}_v\xi_1, \mathcal{K}_v\xi_2\rangle_q.$$

- The Hamiltonian vector field on N is X, defined by $i_X\omega = -dE$. The energy function is invariant under the Hamiltonian flow. In fact,

$$XE = i_X dE = -\omega(X,X) = 0.$$

We have $d\pi_v X = v$. In fact, let $q = \pi(v)$ and $w \in T_q M$ be arbitrary. Define $\xi := \mathcal{I}_v w$. Then, from the expression of ω given above in terms of the Riemannian metric on M,

$$i_X\omega(\xi) = -\langle \mathcal{K}_v\xi, d\pi_v X\rangle = -\langle w, d\pi_v X\rangle$$

and it is a simple check to verify that $-dE(\xi) = \langle v, w\rangle$, which we obtain by recalling that $\mathcal{I}_v w$ is represented by the curve in N given by $\xi(t) = v + tw$. We also have

$$\mathcal{K}_v X = -\mathcal{I}_v \mathrm{grad}_q U.$$

This is essentially Newton's equation of motion. It follows that $X = \mathcal{Z} - (\mathrm{grad}\, U)^V$, where $\mathcal{Z}$ is the canonical vector field and the superscript V denotes vertical lift.

- $\mathcal{L}_Z$, for a vector field Z, denotes the Lie derivative along Z. Recall that the Lie derivative applied to a differential form α satisfies $\mathcal{L}_Z\alpha = di_Z\alpha + i_Z d\alpha$. The Lie derivative of θ and ω along the Hamiltonian vector field X are $\mathcal{L}_X\theta = dL$ and $\mathcal{L}_X\omega = 0$. The first identity is a consequence of $i_X d\theta = -dE$ and of

$$(i_X\theta)_v = \langle v, d\pi_v X\rangle_q = \|v\|_q^2 = E(q, v) + L(q, v),$$

from which we obtain $di_X\theta = dE + dL$. The second identity now results from $d\mathcal{L}_X = \mathcal{L}_X d$.

Proposition 5.1. Let the dimension of M be n. For Newtonian systems $\Omega := \frac{1}{n!}\omega\wedge\cdots\wedge\omega = \frac{1}{n!}\omega^n$ is nowhere vanishing and invariant under the Hamiltonian flow.

Proof. That Ω is nowhere vanishing is due to the Newtonian Lagrangian being regular. Invariance follows from $\mathcal{L}_X\omega = 0$ and since the Lie derivative is an algebraic derivation with respect to the wedge product. □

Definition 5.1 (Liouville measure)**.** The measure obtained by integration with respect to Ω is called the *Liouville measure*. We will at times denote it by $|\Omega|$. In canonical coordinates,

$$|\Omega| = dq_1 \dots dq_n dp_1 \dots dp_n.$$

Invariance of the energy function E under the Hamiltonian flow means that the Hamiltonian vector field X is tangent to the level sets of E. Let us fix a value $\mathcal{E}$ of E and denote by $N(\mathcal{E})$ the level set for $\mathcal{E}$. We assume that $\mathcal{E}$ is a regular value so that $N(\mathcal{E})$ is a smooth submanifold of N, possibly with boundary. The boundary of $N(\mathcal{E})$ consists of the v that project under π to a boundary point of M. We have introduced before the unit normal vector $\eta := \mathrm{grad}\, E/\|\mathrm{grad}\, E\|$ to the level sets of E.

Proposition 5.2. Let X be the Hamiltonian vector field for the energy function E, and η the unit normal vector to the level sets of E defined above. Let $\Omega = \frac{1}{n!}\omega^n$ be the invariant volume form on N and define $\Omega_E := i_\eta\Omega$. Then the following hold.

1. $\Omega = dE \wedge \Omega_E$;
2. $i_X\Omega_E = \frac{1}{(n-1)!}\omega^{n-1} + \frac{1}{(n-2)!}(i_\eta\omega)\wedge dE\wedge\omega^{n-1}$;

3. $\mathcal{L}_X\Omega_E = dE\wedge(i_\eta\mathcal{L}_X\Omega_E)$.

It follows in particular that the restriction of Ω_E to each regular level set $N(\mathcal{E})$ is non-vanishing and invariant under the Hamiltonian flow. Furthermore, the restriction of $i_X\Omega_E$ to the same level sets equals $\frac{1}{(n-1)!}\omega^{n-1}$.

Proof. For part (1), write $dE\wedge\Omega_E = f\Omega$ and take the interior product on both sides with η to conclude that $f\Omega_E = \Omega_E$. As Ω_E does not vanish we must have $f = 1$. For part (2), observe that that

$$i_X i_\eta\omega^n = i_X\left(n(i_\eta\omega)\wedge\omega^{n-1}\right) = n\omega(\eta,X)\omega^{n-1} - n(i_\eta\omega)\wedge(i_X\omega^{n-1}).$$

But $\omega(\eta,X) = dE(\eta) = 1$ by the definition of X and η. Also $i_X\omega^{n-1} = -(n-1)dE\wedge\omega^{n-2}$. For part (3),

$$0 = \mathcal{L}_X\Omega = \mathcal{L}_X(dE\wedge\Omega_E) = dE\wedge\mathcal{L}_X\Omega_E = 0,$$

then apply interior multiplication on both sides of the equality by η. □

In classical statistical physics, when considering and isolated macroscopic mechanical systems consisting of a large number of microscopic particles, a postulate due to Gibbs asserts that the equilibrium distribution of states of the system for a given total energy $\mathcal{E}$ is the uniform distribution relative to Ω_E on the energy level set $N(\mathcal{E})$. This postulate is known by the name *Gibbs microcanonical ensemble*. Assigning probabilities to events, regarded as Borel subsets of $N(\mathcal{E})$, using the measure $|\Omega_E|$ only makes sense, naturally, when the total $|\Omega_E|$-volume of a given energy level set is finite.

Proposition 5.3. Let D be a two-dimensional submanifold of $N(\mathcal{E})$ diffeomorphic to a closed disc with smooth boundary D, and $T: D\to\mathbb{R}$ a smooth function. Define $\Psi(v) := \Phi_{T(v)}(v)$, where Φ_t is the Hamiltonian flow, and let D' be the image of D under Ψ. Then $\Psi^*\omega = \omega$.

Proof. Let $\mathcal{T}$ denote the 2-dimensional set consisting of all $\Phi_t(v)$ for $v\in\partial D$ and $t\in[0,T(v)]$. We will call $\mathcal{T}$ a *flow tube*. Note that $\int_{\mathcal{T}}\omega = 0$. In fact, at every $v\in\mathcal{T}$, the tangent space to $\mathcal{T}$ at v is spanned by X_v and another vector ξ which is, naturally, tangent to $N(\mathcal{E})$. It follows that $\omega_v(X_v,\xi) = -dE(\xi) = 0$, hence $\omega = 0$ on $\mathcal{T}$. The boundary of $\mathcal{T}$ is the union of ∂D and $-\Psi(\partial D)$ where the negative sign indicates orientation. Therefore, recalling that $\omega = d\Theta$,

$$0 = \int_{\mathcal{T}}\omega = \int_{\partial\mathcal{T}}\Theta = \int_{\partial D}\theta - \int_{\Psi(\partial D)}\theta = \int_{\partial D}(\Theta - \Psi^*\Theta) = \int_D(d\Theta - d\Psi^*\Theta) = \int_D(\omega - \Psi^*\omega).$$

Since the same argument applies to an arbitrary disc in D, we conclude that $\omega = \Psi * \omega$. That is, the pull-back of ω on $\Psi(D)$ equals ω on D. □

Recall, from Corollary 3.3 that $\omega = d\Theta$ is a symplectic form on sets of the form $N_S(\mathcal{E})\setminus TS$. So it makes sense to define a measure $|\omega^{n-1}|$ on this set.

Proposition 5.4. Let S be a smooth hypersurface in M and $\mathcal{E}$ a regular value of the energy function, with level set $N(\mathcal{E})$. Let $N_S(\mathcal{E})$ be the subset of $N(\mathcal{E})$ projecting into S under the base point map. Let $U = N_S(\mathcal{E})\setminus TS$ and define $T(v) = \inf\{t>0 : \Phi_t(v)\in N_S(\mathcal{E})\}$ for each $v\in U$, where $T(v) = +\infty$ if the set of which we are taking the infimum is empty. Let Q be the subset of U in which T is finite and define $\Psi(v) := \Phi_{T(v)}(v)$ for $v\in Q$. Then the measure $\mu := |\omega^{n-1}|$ is invariant under Ψ; that is, $\mu(A) = \mu(\Psi^{-1}(A))$ for any Borel set $A\subset Q$.

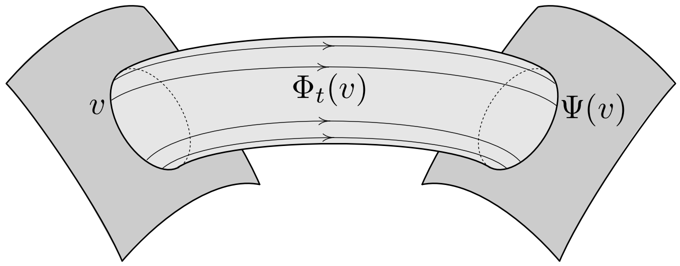

Figure 8: A flow tube used in the proof of Proposition 4.3

Proof. * □

Proposition 5.5. Let S be an oriented smooth hypersurface in M with orientation defined by a unit normal vector field $n(q)$. At each $q \in S$ we define the reflection map $R_q(v) := v - 2\langle v, n\rangle_q n$. Note that R defines a diffeomorphism from $N_S(\mathcal{E})$ to itself. Then $R^*\Theta = \Theta$.

Proof. Note that $\pi \circ R = \pi$ and observe that$\langle v, d\pi_v\xi\rangle_q = \langle R_q(v), d\pi_v\xi\rangle_q$ for all ξ tangent to $N_S(\mathcal{E})$ at v since $d\pi_v\xi$ is tangent to S. It follows that

$$(R^*\Theta)_v(\xi) = \Theta_{R(v)}(dR_v\xi) = \langle R(v), d(\pi \circ R)_v\xi\rangle_q = \langle R(v), d\pi_v\xi\rangle_q = \langle v, d\pi_v\xi\rangle_q = \Theta_v(\xi).$$

This shows that Θ is invariant under R. □

We would like to obtain a more concrete description of the invariant volume on $N_S(\mathcal{E}) \setminus TS$ for a hypersurface S in M in terms of the Riemannian volume form on M and the standard volume on spheres in T_qM. We do this next. Recall that ω defines a symplectic form on the $(2n-2)$-dimensional space $N_S(\mathcal{E}) \setminus TS$, where S is some hypersurface in the n-dimensional manifold M. We assume that S is oriented, and let $n(q)$, $q \in S$, be a vector unit length perpendicular to T_qS for each q. Then $N_S(\mathcal{E}) \setminus TS$ is a disjoint union of the two sets

$$N_S^{\pm}(\mathcal{E}) := \{(q, v) \in N_S(\mathcal{E}) : \pm\langle v, n\rangle_q > 0\}.$$

Note that on a neighborhood of any point q where $U(q) < \mathcal{E}$, the set $N_S^{\pm}(\mathcal{E})$ is diffeomorphic to a product of an open set in S and the open hemisphere $\{x = (x_1, \dots, x_n) \in S^{n-1} : \pm x_n > 0\}$. Our immediate goal is to express the invariant volume form on $N_S^{\pm}(\mathcal{E})$ in terms of the Riemannian volume form on S and the volume form on hemispheres. We will do this for $N_S^{+}(\mathcal{E})$, the negative case being similar.

Also note that, for any given value of $\mathcal{E}$ the only allowable configurations are those q for which $\mathcal{E} - U(q) \geq 0$ because this difference is the (non-negative) kinetic energy. Therefore, it makes sense to define the function

$$h_{\mathcal{E}}(q) := \sqrt{2(\mathcal{E} - U(q))}$$

on the invariant set $M(\mathcal{E}) := \{q \in M : U(q) \leq \mathcal{E}\}$. Now let W be any open set in $M(\mathcal{E})$ on which the tangent bundle of M is trivial, so that it is possible to define on W a smooth family of orthonormal vector fields $e_1, \ldots, e_n$. We do this in such a way that on $W \cap S$, if this set is non-empty, e_n is the unit normal vector denoted n above. We denote by $N_W^+(\mathcal{E})$ the subset of $N(\mathcal{E})$ projecting to W under the base-point map $\pi : N \to M$, by S^{n-1} the unit sphere in $\mathbb{R}^n$, and by $F_{\mathcal{E}} : W \times S^{n-1} \to N_W(\mathcal{E})$ the map defined by

$$F_{\mathcal{E}}(q, u) := \left(q, h_{\mathcal{E}}(q) \sum_{i=1}^{n} u_i e_i(q)\right).$$

We call $F_{\mathcal{E}}$ a *frame map*. The Riemannian volume form on M is the n-form ω^M (defined up to a sign) such that $\omega_q^M(e_1, \ldots, e_n) = 1$ for any orthonormal basis $\{e_1, \ldots, e_n\}$ at q. Similarly, define ω^S, for a hypersurface S in M. Let the volume form on the sphere S^{n-1} be the $(n-1)$-form denoted by ω^{sphere}. Recall the forms Ω_E on $N(\mathcal{E})$ and $\Omega_E^S := i_X \Omega_S$ on $N_S(\mathcal{E})$ of degrees $2n-1$ and $2n-2$, respectively.

Theorem 5.1 (Invariant volume measures)**.** For any choice of orthonormal frame over an open set $W \subset M(\mathcal{E})$ and given the frame map $F_{\mathcal{E}}$ defined above we have

$$F_{\mathcal{E}}^* \Omega_E = \pm n h_{\mathcal{E}}^{n-2} \omega^M \wedge \omega^{\text{sphere}}.$$

If W is a neighborhood of a point in S, we similarly have

$$F_{\mathcal{E}}^* \Omega_E^S = \pm \cos\theta\, h_{\mathcal{E}}^{n-1} \omega^S \wedge \omega^{\text{sphere}}$$

where $\theta(v)$ is the angle that $v \in T_q M$ makes with the normal vector $n(q)$ for $q \in S$. Apart from the unspecified signs, these expression do not depend on the choice of local orthonormal frames.

Proof. * □

It is helpful to give another proof of the theorem in the special, and much simpler, case of of an n-dimensional billiard system in Euclidean space with 0 potential function. We use the same notation as in Exercise 3.6, where M is an n-dimensional submanifold of $\mathbb{R}^n$ with smooth boundary and each element of TN is written as $(q, v, \dot{q}, \dot{v})$.

Exercise 5.1. Let $M = \mathbb{R}^n$ with Lagrangian $L(q, v) = \frac{1}{2}|v|^2$, where $|v|^2 = v \cdot v$. Recall the vector field η from Exercise 3.6. Show the following.

1. $\Omega = \frac{1}{n!}\omega^n = (-1)^{\frac{n(n-1)}{2}} dq_1 \wedge \cdots \wedge dq_n \wedge dv_1 \wedge \cdots \wedge dv_n$
2. The interior product of ω^n by η is $i_\eta \omega^n = \frac{n}{|v|} \Theta \wedge \omega^{n-1}$.

Exercise 5.2. Let $S^{n-1} = \{v \in \mathbb{R}^n : |v| = 1\}$ be the unit sphere in $\mathbb{R}^n$. The volume form on S^{n-1} is

$$\omega^{\text{sphere}} = i_v(dv_1 \wedge \cdots \wedge dv_n) = \sum_{j=1}^{n} (-1)^j v_j dv_1 \wedge \cdots \wedge dv_{j-1} \wedge dv_{j+1} \wedge \cdots \wedge dv_n$$

where i_v indicates interior product with the vector field $\sum_j v_j \frac{\partial}{\partial v_j}$.

Exercise 5.3. Show that the form Ω_E is

$$\Omega_E = \frac{1}{(n-1)!}\frac{1}{|v|}\Theta\wedge\omega^{n-1} = (-1)^{\frac{n(n-1)}{2}}\frac{1}{|v|}dq_1\wedge\cdots\wedge dq_n\wedge\omega^{\text{sphere}}.$$

Let us now describe the invariant volume form on the space of boundary states with energy equal to $1/2$. That is, the set of unit vectors in N with base point in the boundary of M. We denote this set by N_S^{unit}, where S stands for the boundary of M. Let $n(q)$ be the inward pointing unit normal vector to S at $q \in S$. The tangent bundle of N_S^{unit} is

(5.1) $$TN_S^{\text{unit}} = \{(q, v, \dot{q}, \dot{v}) \in TN : q \in S, v\cdot\dot{v} = 0, \dot{q}\cdot n(q) = 0\}.$$

Let $e_1,\dots,e_n$ be any orthonormal basis of $\mathbb{R}^n$, with dual basis denoted $e_1^*,\dots,e_n^*$. The dual basis is defined by the property $e_i^*(e_j) = \delta_{ij}$ where $e_i^*(e_j)$ indicates the pairing of vector and covector. Under the identification of the Horizontal subspaces in TN with tangent spaces TM, the Hamiltonian vector field $X_{(q,v)} = (q, v, v, 0)$ may be written in terms of the basis $\{e_j\}$ as $X_{(q,v)} = \sum e_j^*(v)e_j$.

Exercise 5.4. Show that, at each boundary state, the form $i_X\Omega_E$ can be expressed in terms of any orthonormal basis $e_1,\dots,e_n$ of $\mathbb{R}^n$ as

$$i_X\Omega_E = (-1)^{\frac{n(n+1)}{2}}\left(\sum_j(-1)^j\frac{v\cdot e_j}{|v|}e_1^*\wedge\cdots\wedge e_{j-1}^*\wedge e_{j+1}^*\wedge\cdots\wedge e_n^*\right)\wedge\omega^{\text{sphere}}.$$

By the description of TN_S^{unit} given in the above 4.1, the form between parenthesis is evaluated on vectors tangent to S. Therefore, if we use an orthonormal basis such that e_n is normal to T_qS, we obtain that this form (at q) reduces to

$$(-1)^n\frac{v\cdot e_n}{|v|}e_1^*\wedge\cdots\wedge e_{n-1}^* = (-1)^n\frac{v\cdot e_n}{|v|}\omega^S$$

where ω^S denotes the Riemannian volume form of S for the Riemannian metric induced on S as a hypersurface of $\mathbb{R}^n$. Assuming that $E = 1/2$ so that $|v| = 1$, and denoting by $n(q)$ the inward pointing unit normal vector to S at q, conclude that

$$(i_X\Omega_E)_{(q,v)} = (-1)^{\frac{n(n+3)}{2}}v\cdot n(q)\omega^S\wedge\omega^{\text{sphere}}.$$

Definition 5.2 (The billiard map)**.** Let M be a Riemannian manifold with boundary S and $L(q,v)$ a Newtonian Lagrangian function on $N = TM$. Denote by $n(q)$ the inward pointing unit normal vector to S at q. Let $N_S \subset N$ consist of the (q,v) such that $q \in S$ and for $\epsilon \in \{+,-\}$ denote by N_S^ϵ the two components of $N_S \setminus TS$, where N_S^ϵ consists of the (q,v) such that $\epsilon\langle v, n_q\rangle > 0$. Define the reflection map $R : N_S^- \to N_S^+$ by

$$R(q,v) = (q, v - 2\langle v, n_q\rangle n_q).$$

Let Φ_t be the Hamiltonian flow on N, and for $(q,v) \in N_S^+$ let $T(q,v) = \inf\{t > 0 : \Phi_t(q,v) \in N_S\}$. We set $T = \infty$ if the set of which we are taking the infimum is empty. Let N_{billiard} be the subset of N_S^+ on which T is finite, and define on N_{billiard} the map $\Psi(q,v) := \Phi_{T(q,v)}(q,v)$. Finally, define the map $B : N_{\text{billiard}}^+ \to N_S^+$ by $B := R\circ\Psi$. We call B the *billiard map* of the Hamiltonian system with boundary. Then μ is invariant under the billiard map B.

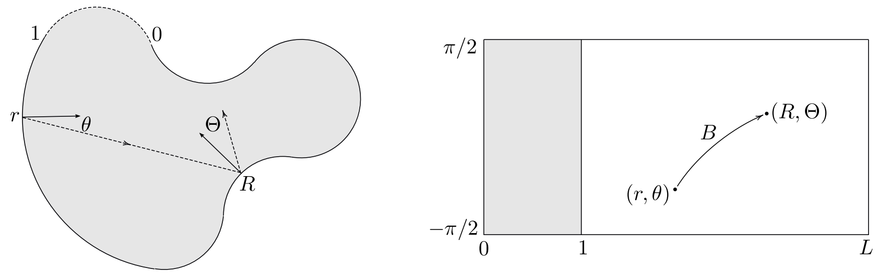

Figure 9:

Corollary 5.1. We use the notations of Definition 4.2 and assume that the Lagrangian function is $L(q,v) = \frac{1}{2}\|v\|_q^2$. Define on N_S^+ the measure μ such that

$$d\mu(q,v) = v \cdot n(q) d\mathrm{Vol}^S(q)\, d\mathrm{Vol}^{\text{sphere}}(v)$$

where $d\mathrm{Vol}^S(q)$ is the Riemannian volume element on the hypersurface S and $d\mathrm{Vol}^{\text{sphere}}(v)$ is the volume element on the unit (hemi-)sphere in the tangent space of M at q. For billiard systems in dimension 2, $d\mathrm{Vol}^S(q)$ is ds, where s is the variable parametrizing the boundary of the table by arclength, and $d\mathrm{Vol}^{\text{sphere}}(v)$ is the angle measure $d\theta$, where $\theta \in (-\pi/2, \pi/2)$ is the angle that a unit vector v at q makes with the inward pointing unit normal vector $n(q)$. Thus if the billiard table has finite perimeter of length L, the probability measure $d\mu(s,\theta) = \frac{1}{2L}\cos\theta\, ds\, d\theta$ defined on the phase space $[0, L] \times (-\pi/2, \pi/2)$ is invariant under the billiard map.

Figure 10 given an interpretation of the quantity $\cos\theta\Delta s$, where $\Delta s = l$ in the figure, for the special case of a polygonal billiard table.

Figure 11 illustrates the invariance of the measure μ under the billiard map in dimension 2. The table is a unit square with a circular scatterer in the middle. More precise, we consider the map that gives the first return to the top side of the square of trajectories that begin at that side. Thus, for each initial condition, the return state is the image of a number of iterations of the billiard map. The initial states indicated in Figure 11 are uniformly distributed over a small rectangle in $[0, 1] \times (-\pi/2, \pi/2)$. The image of this small rectangle under the return map is widely dispersed over the phase space, but has the same measure $d\mu = \frac{1}{2}\cos\theta\, ds\, d\theta$. Note that $L = 1$ here since we are restricting attention only to the top side of the square.

Another interesting interpretation of the billiard measure in dimension 2 is shown in Figure 13. The phase space of the billiard system is in this case taken to be the rectangle $[0, 1] \times (0, \pi)$, where the angle θ is measure from the positive tangent direction to the boundary of the table, oriented counterclockwise, and the total perimeter is assumed to have length 1. As indicated in the figure, there is a natural bijection between the phase space and the 2-sphere minus the north and south poles. Regarding the sphere as the phase space of the billiard system, the invariant billiard measure μ turns out to be the normalized area measure of the sphere. In other words, the billiard map corresponds to an area preserving map of the 2-sphere.

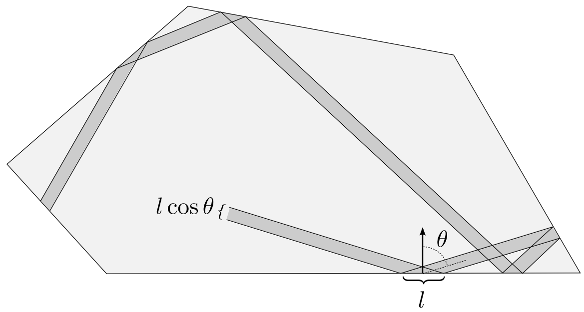

Figure 10: Illustrating the invariant billiard measure in dimension two for a polygonal billiard table. A long tape of uniform width is folded into a beam of billiard trajectories. At each collision the length l along the fold of the tape times the angle θ the tape makes with the normal vector at the collision point is constant, equal to the width. So the width is a conserved quantity of the billiard motion.

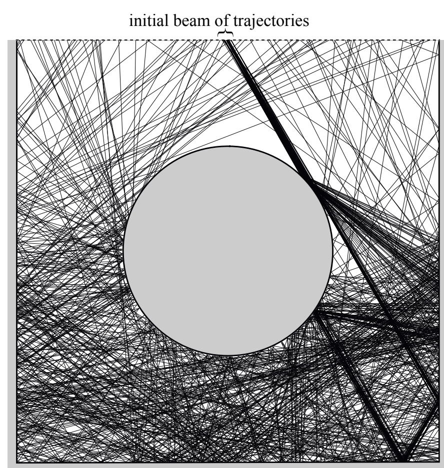

Figure 11: A beam of 10000 billiard trajectories are traced with initial position on the dashed line segment, identified with the interval $[0,1]$, on top of the rectangular table. They are stopped at their first return to that side. The initial conditions (x,θ) are random points uniformly distributed on a phase space rectangle with $0.5-0.01 \le x \le 0.5+0.01$ and $-\pi/6-0.01 \le \theta \le -\pi/6+0.01$. The image of this small rectangle under the first return billiard map is shown in Figure 12.

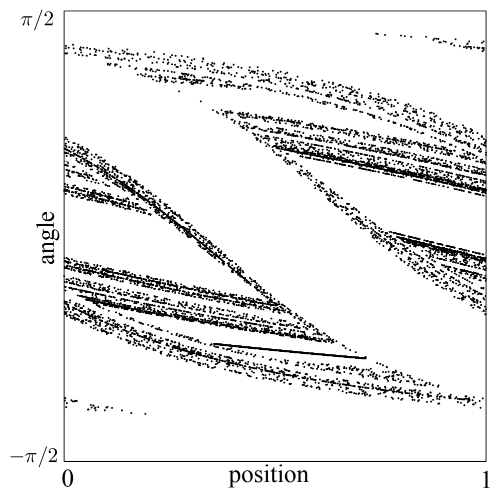

Figure 12: This scattered plot shows the end states of the first return to the dashed line segment of Figure 11 of the 10000 trajectories. The region occupied by the dots roughly represents the image under the first return map of the small rectangle in phase space of initial conditions described in the previous figures. Although the initial rectangle is greatly distorted by the map, their measures with respect to $d\mu = \frac{1}{2}\cos\theta \, ds \, d\theta$ are the same.

6 Ergodic theory

Ergodic theory deals with measure theoretic issues of dynamical systems. For us, It is the natural place to begin to introduce probabilistic ideas in our study of mechanical systems. To begin to see the types of questions considered in ergodic theory, let us explore a very concrete problem related to first return to a subset of the boundary of a billiard system.

Consider the billiard system shown in Figure 14. We think of it as a sort of two-dimensional cave whose piecewise smooth inner walls are perfectly reflecting mirrors. The cave is open at the top; we assume that the mouth of the cave is an interval of length 1, which we identify as the unit interval $[0,1]$. From a point r in $[0,1]$ we flash into the cave at an angle $-\pi/2 < \theta < \pi/2$ an infinitely narrow light beam, which we think of as a single light ray. We then wait for the light ray to reemerge from the cave and register its position and angle as it leaves the cave. We may repeat this experiment as many times as we want. Here are a few questions of possible interest.

- How long, on average, does it take for the light ray to reemerge?
- How many times, on average, does a light ray bounce off the inside walls?
- What is the likely angle at which the light ray will exit the cave?
- How do these quantities depend on the internal shape of the cave?

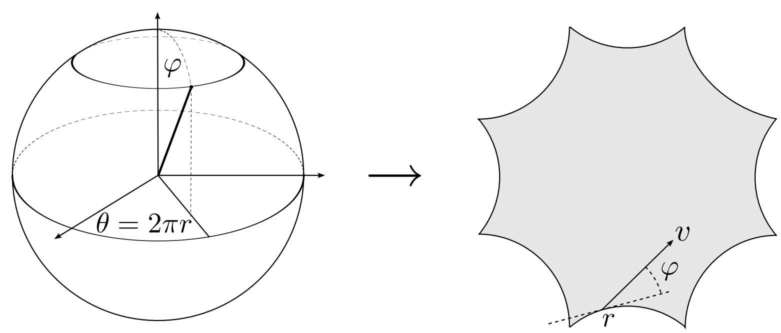

Figure 13: Assuming that the billiard table has perimeter 1, parametrized by arc-length parameter $r \in [0,1]$, and letting $\varphi \in (0,\pi)$ denote the angle that v makes with the tangent line to the boundary of the table, then the phase space $[0,1]\times(0,\pi)$ of the billiard system can be identified with the two-dimensional sphere (minus the north and south poles) via spherical coordinates. Under this identification the invariant probability billiard measure $\frac{1}{2}\sin\varphi\, dr\, d\varphi$ corresponds to the normalized area measure $\frac{1}{4\pi}\sin\varphi\, d\theta\, d\varphi$ of the unit two-sphere.

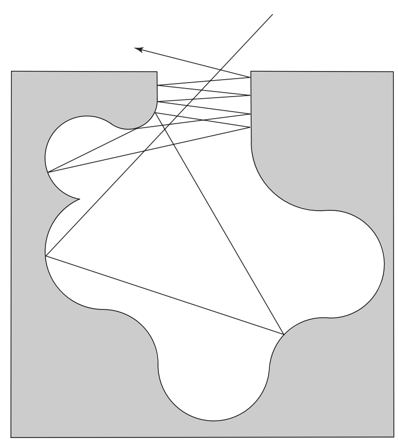

Figure 14: Light rays are flashed into a two-dimensional cave with piecewise smooth perfectly reflecting inner walls. We assume that the total area inside the cave is finite.

For these questions to make sense, we need to address the following issues.

- What do we mean by "on average"?
- How do we know that light rays will reemerge in the first place?
- If light rays can get trapped inside, how likely is it to happen?
- Do we have any right to expect simple answers to these questions?

6.1 Trapped trajectories

A natural first question to investigate is whether it is possible for trajectories starting outside the cave can get trapped inside and never again reemerge. We consider this problem for the specific example depicted in Figure 15.

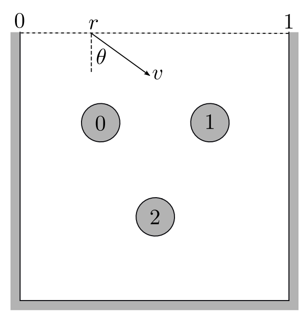

Figure 15: A cave with circular scattering walls inside. It is possible for uncountably many light trajectories starting outside the cave to get trapped inside.

Exercise 6.1. Show that there are uncountably many billiard trajectories for the system described in Figure 15 that begin outside the square and become trapped inside. In fact, given any infinite sequence $a_1, a_2, a_3, \ldots$, where $a_j \in \{0,1,2\}$ such that $a_{j+1} \neq a_j$ for all j, show that there is a trajectory that bounces off the circular scatterers in the order set by the given sequence, never touching the straight sides of the table.

7 Systems of billiard-type

Although there may be plenty of trapped trajectories, the questions asked at the beginning of this section are still meaningful because, it turns out, the probability of trapping is aways zero, so long as the area of the billiard table is finite.

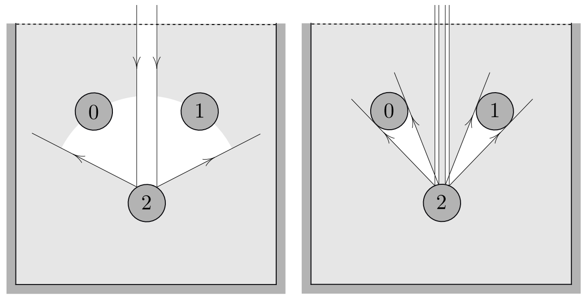

Figure 16: First two steps in the construction of a Cantor set of trapped trajectories.

References

[1] M. do Carmo, *Riemannian Geometry*, Birkhäuser, 1993.

[2] G.B. Folland, *Real Analysis*, Wiley, 1999.

[3] P. Gaspard, *Chaos, scattering and statistical mechanics*, Cambridge University Press, 1998.

[4] L.C. Grove, C.T. Benson, *Finite Reflection Groups*, Springer,Graduate Texts in Mathematics, 99, 1985.

[5] J.M. Lee, *Introduction to Smooth Manifolds*, Springer Graduate Texts in Mathematics, 218, 2003.

[6] H. Masur and S. Tabachnikov, *Rational Billiards and Flat Surfaces*.