NOTES FOR THE CIMPA-NIMS RESEARCH SCHOOL DAEJEON, 25.5.-3.6. 2015

FELIX SCHLENK

1. Configuration spaces and linkages

1.1. **Configuration spaces and phase spaces.** The set of configurations Q that a given dynamical system can attain is called the *configuration space* of the system. This is the "space of candidates" for the positions that the system may attain, and given an initial condition, the task is to find out which position the system actually will attain at a given time.

We shall usually assume that the set Q is a smooth manifold, so that tools from analysis can be used to study the dynamics. Often, but not always, we shall assume that Q is compact (under natural assumptions on the dynamical system, this guarantees that the motion exists for all times).

In this section we first define the tangent bundle and the cotangent bundle of a configuration space, and then give many examples of configuration spaces of concrete dynamical systems. Most of these spaces will be compact, but some are not.

1.2. **Tangent space and cotangent space.** If we just worked on the space Q, then the set of the trajectories of our dynamics on Q does not fit together to form a flow: In a given point $q \in Q$, one can move in various directions, and hence the point $q \in Q$ and a time $t_0 \in \mathbb{R}$ would not give rise to an initial value problem. We therefore rather look at the points $q \in Q$ *and*, for each q, at all the possible vectors v at q into which the point q may move, that is, the tangent space T_qQ of Q at q. The union of these tangent spaces,

$$TQ = \bigcup_{q \in Q} T_qQ$$

is the *tangent bundle* of Q. The manifold structure of Q gives rise to a manifold structure of TQ in a natural way: If $\varphi\colon U \to Q$ is a chart with $U \subset \mathbb{R}^n$, then

$$T\varphi\colon TU = U \times \mathbb{R}^n \to TQ, \quad T\varphi(x,y) = (\varphi(x), D\varphi(x)y)$$

is a chart for TQ, and if $\{(U_j, \varphi_j)\}$ is an atlas for Q, then $\{(TU_j, T\varphi_j)\}$ is an atlas for TQ.

Date: May 14, 2015.

Examples of configuration spaces. The following example shows that the concept of configuration space is non-trivial, and can be surprisingly helpful.

Example 1.1. Two nonintersecting roads (of zero width) lead from town A to town B. It is known that two (pointlike) cars, connected by a rope of length $2r$, were able to travel on these different roads from A to B without breaking the rope. Is it possible for two circular wagons of radius $> r$, one starting at A and one at B, whose centers move along these roads, to reach the opposite town without colliding?

We invite the reader to guess the answer intuitively, and to compare (in case of a correct answer) his intuition with the following argument:

...

It is rather clear that these two curves must intersect.

Exercise 1.2. Give a mathematical argument for the fact that the curves γ_1, γ_2 must intersect.

1.1. The pendulum. The configuration space is thus the circle S^1.

1.2. The double pendulum. The configuration space is thus the 2-torus $T^2 = S^1 \times S^1$.

1.3. The spherical pendulum. The configuration space is thus the 2-sphere S^2. We agree that the presence of the mounting makes it hard to realize all of S^2 as configuration space. However, you may think of the linkage rather as a 2-atom molecule, with one atom at rest, and the rod replaced by a force.

Coupled penduli.

The configuration space of two mechanical systems, modeled on Q_1 and Q_2, is the product $Q_1 \times Q_2$. If the dynamics do not interact, then we will just have a product situation for the dynamics too, and may equally well study the two dynamics separately. If the systems are coupled, however, then the dynamics may be far from split.

Example coupled penduli spring, Bild

2. The rigid body. A rigid body K is a body that is rigid, i.e., a system of point masses in $\mathbb{R}^3$, constrained by holonomic relations expressed by the fact that the distance between points is constant:

$$|x_i - x_j| = r_{ij} = \text{const}$$

Approximate examples are tops, airplanes (hopefully), and planets.

Exercise 1.3. Show that the configuration space Q of K is

$\mathbb{R}^3$ if K is a point;

$\mathbb{R}^3 \times S^2$ if K is not a point and is contained in a line;

$\mathbb{R}^3 \times \mathbf{SO}(3)$ if K is not contained in a line. Here, $\mathbf{SO}(3)$ is the group of rotations of $\mathbb{R}^3$.

Proof. Refer to [1, Ch. 28], or (a bit less precise): A translation by $-\mathcal{O} \in \mathbb{R}^3$ moves the center of mass $\mathcal{O}$ of K to the origin of $\mathbb{R}^3$. If K is contained in a line, then this oriented line is determined by a point on S^2. Otherwise, let $F = \{e_1, e_2, e_3\}$ be the standard orthonormal frame in $\mathbb{R}^3$. Then any other position of K relative to the origin corresponds to a unique other orthonormal frame F'. There is a unique rotation taking F to F'. $\square$

These configuration spaces are not compact. However, the description of the motion of K decomposes into two parts: The motion of its center of mass $\mathcal{O}$ in $\mathbb{R}^3$, and the motion of K relative to $\mathcal{O}$. The first problem is sometimes easily solved. For instance, the center of mass of a free rigid body moves with constant speed along a line. One may then restrict attention to the motion of K around $\mathcal{O}$, on the reduced configuration space S^2 or $\mathbf{SO}(3)$.

Exercise 1.4. Show that the space $\mathbf{SO}(3)$ is diffeomorphic to real projective space $\mathbb{R}\mathrm{P}^3 = S^3/\{x \sim -x\}$.

The above examples suggest that closed manifolds Q that appear as configuration spaces of a classical mechanical system are (quotients of) spheres and products of such manifolds. It may therefore look awkward to develop a theory for arbitrary closed manifolds Q. As the following examples show, manifolds more complicated than spheres naturally do arise already in simple mechanical devices.

4. Planar linkages. A planar linkage is a system of rigid rods with hinge connections. Some of the hinges are fixed, and some can freely move. As before, we shall paint the fixed hinges in black and the movable ones as circles. We have already met the two most simple linkages: A pendulum has one rod, with one end fixed. A double pendulum consists of two rods, fixed at one end point.

As a warm-up, we look at the linkage drawn in Figure 1, and determine the configuration spaces $Q(d)$ for each $d \geq 0$.

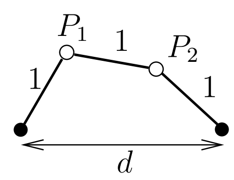

FIGURE 1. A linkage with three rods.

Consider the possible positions of the pin P_1. For $0 \leq d \leq 1$, the pin P_1 can move along a circle, while for $d \in (1, 3)$ it can move along a closed segment. For each position of P_1 the other pin P_2 can attain one or two positions. For $0 \leq d < 1$, we have $Q(d) = S^1 \amalg S^1$. When d tends to 1, the left points of the two circles merge, and at the right a new circle appears. For $d \in (1, 3)$, $Q(d)$ is a circle. And of course, $Q(3)$ is a point and $Q(d)$ is empty for $d > 3$. Notice that the space $Q(d)$ does not change in a continuous way at $d = 1$, and that $Q(1)$ is not a manifold.

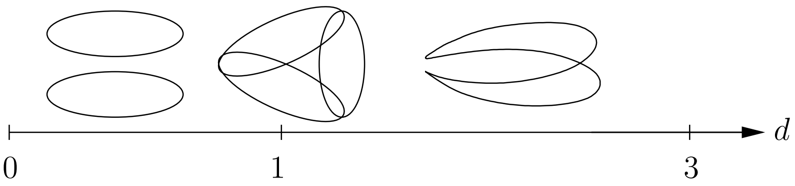

FIGURE 2. The configuration spaces $Q(d)$.

The triple arm linkage. Consider now the triple-arm linkage drawn in Figure 3: At each vertex of an equilateral triangle of side length ℓ, fix an arm made of two rods of lengths r and s, where $r > s$ and $r + 2s > \ell$, and join the three arms at their moving ends. In order to understand the configuration space Q, we first draw the set of possible positions of the central pin O. Each arm keeps the pin within a ring centered at its anchor, of inner radius $r - s$ and outer radius $r + s$. Hence the pin O can reach any point that lies in the intersection of the three rings corresponding to the three arms. Since $r > s$ and $r+2s > \ell$, this region is a curvilinear hexagon, as drawn in Figure 3.

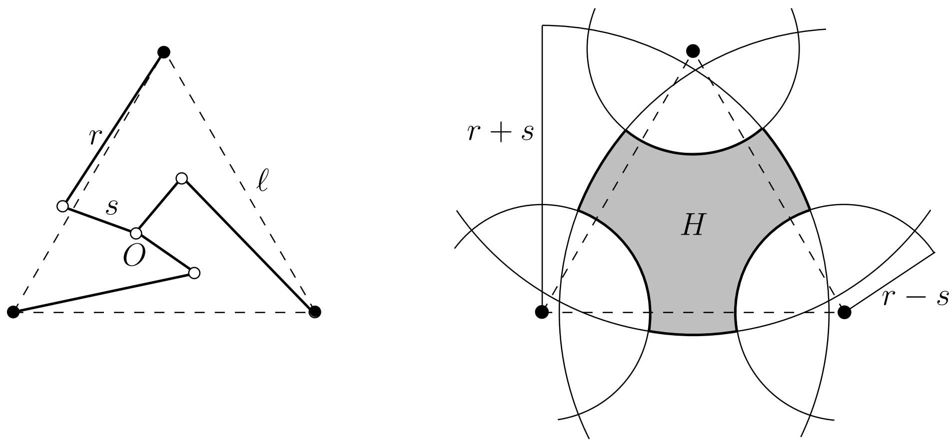

FIGURE 3. The triple arms linkage, and the hexagon H.

For each point O in the interior of the hexagon, the elbow of each arm can be bent in one of two ways (to the left or to the right). The total number of configurations for each point in the interior is therefore $2^3 = 8$. We conclude that Q is obtained from eight copies of the hexagon, each hexagon corresponding to a different combination of elbows (right-right-right, left-right-right, etc.). Since Q is clearly connected, these eight pieces are to be glued together. Let's see how.

When O lies on an edge of the hexagon (but not on a vertex), exactly one arm is fully straight or fully folded while the other two arms are not aligned, and hence there correspond $2^2 = 4$ configurations to this position of O. Given such a position, we can reach nearby positions in which the previously straight arm is now slightly bent either to the left or to the right, while the other two arms are still bent in the same direction. It follows that from such a position we can reach exactly two hexagons. In other words, at each edge of Q exactly two hexagons meet.

When O lies on a vertex of the hexagon, exactly two arms are fully straight or fully folded, hence there correspond $2^1 = 2$ configurations to this position of O. This time, given such a position, we can reach nearby positions in which the two previously straight arms are now slightly bent to the right-right or right-left or left-right or left-left, while the other arm is still bent in the same direction. It follows that from such a position we can reach exactly four hexagons. In other words, at each vertex of Q exactly four hexagonal faces meet.

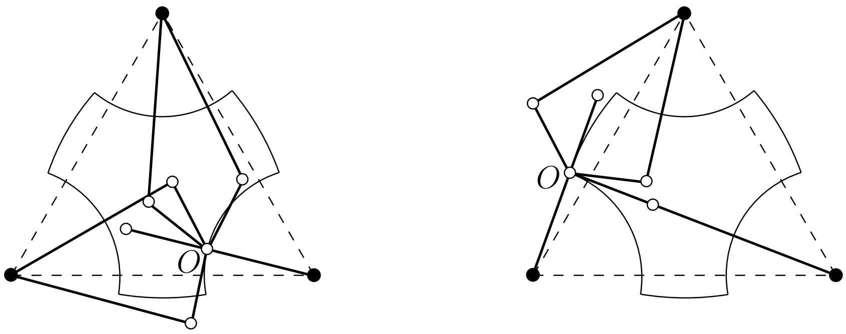

FIGURE 4. When O lies on an edge, or on a vertex, of the hexagon.

We are now ready to compute the number of faces F, the number of edges E, and the number of vertices V of our polyhedral decomposition of Q obtained by the above gluing of the eight hexagons: Of course, $F = 8$. The 8 hexagons have 8×6 edges in total. Since they are glued in pairs, $E = 24$. Finally, there are 8×6 vertices in total, but each vertex belongs to four faces, hence $V = 12$. It follows that the Euler characteristic of the closed surface Q is $\chi(Q) = F - E + V = 8 - 24 + 12 = -4$. Therefore, Q is either a closed orientable surface Σ_3 of genus 3, or the non-orientable closed surface obtained as connected sum of three Klein bottles.

Exercise 1.5. Show that Q is orientable. Hence Q is diffeomorphic to Σ_3.

Exercise 1.6. Show that the configuration space of the linkage in Figure 5 is the closed orientable surface Σ_2 of genus 2.

Exercise 1.7. Consider the linkage drawn in Figure **??**: The two ends are fixed at distance d, the three hinges are free, and each rod has length 1. Denote by $Q(d)$ the configuration space of this linkage. Of course, $Q(d)$ is empty for $d > 4$ and $Q(d)$ is point for $d = 4$.

(i) $Q(3.9)$ is the sphere S^2.

(ii) $Q(1)$ is the orientable surface of genus 4.

The exercise shows that $Q(d)$ is not a smooth surface for all $d \in [1, 3.9]$: As d moves from 1 to 3.9, the topology of the configuration space changes at some values of d. One

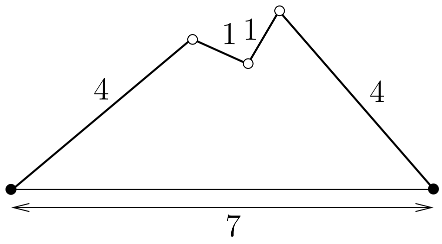

FIGURE 5. A linkage with $Q = \Sigma_2$.

can show that given a linkage, for a generic set of lengths of the arms the configuration space is a smooth manifold, [7,].

We have now found planar linkages that have Σ_g as configuration space for $g = 0, 1, 2, 3, 4$. Surprisingly, any closed manifold can be realized as (part of) the configuration space of a planar linkage:

Theorem 1.8. *Let M be a closed smooth manifold. Then there exists a planar linkage whose configuration space is diffeomorphic to a finite disjoint union of copies of M.*

Exercise 1.9. Let Σ be a closed surface. A *triangulation* of Σ is a decomposition of Σ into triangles "which fit together nicely". (What do we mean exactly? Consult [11, p. 16] if necessary.) It is quite believable, and true, that Σ can be triangulated. Given a triangulation $\mathcal{T}$ of Σ, define $\chi(\Sigma; \mathcal{T})$ as the alternating sum $V - E + F$, where V, E, F are the number of vertices, edges, faces of $\mathcal{T}$.

(i) Prove that $\chi(\Sigma; \mathcal{T})$ does not depend on the triangulation $\mathcal{T}$. The common value $\chi(\Sigma)$ is the *Euler characteristic* of Σ.

Hint: Use the quite believable, and true, fact that two triangulations of Σ have a common refinement.

(ii) Derive Euler's polyhedron formula: For any polyhedral decomposition of Σ, the number of faces minus the number of edges plus the number of vertices equals $\chi(\Sigma)$.

(iii) Compute $\chi(S^2) = 2$ and $\chi(T^2) = 0$.

(iv) Given two closed surfaces Σ and Σ', consider their connected sum $\Sigma \# \Sigma'$. Prove that $\chi(\Sigma \# \Sigma') = \chi(\Sigma) + \chi(\Sigma') - 2$.

(v) Conclude that for the oriented surface of genus g,
$$\chi(\Sigma_g) \;=\; 2 - 2g.$$

(vi) Every non-orientable closed surface Σ has an orientable double cover $\widehat{\Sigma}$. Prove that $\chi(\widehat{\Sigma}) = 2\chi(\Sigma)$.

2. Closed geodesics on Riemannian manifolds

We cannot do without citing the following famous lines of Poincaré's despite its somewhat martial language.

Poincaré 1905 [13]

> *Dans mes* Méthodes nouvelles de la Mécanique céleste *j'ai étudié les particularités des solutions du problème des trois corps et en particulier des solutions périodiques et asymptotiques. Il suffit de se reporter à ce que j'ai écrit à ce sujet pour comprendre l'extrême complexité de ce problème ; à coté de la difficulté principale, de celle qui tient au fond même des choses, il y a une foule de difficultés secondaires qui viennent encore compliquer la tâche du chercheur. Il y aurait donc intérêt à étudier d'abord un problème où on rencontrerait cette difficulté principale, mais où on serait affranchi de toutes les difficultés secondaires. Ce problème est tout trouvé, c'est celui des lignes géodésiques d'une surface ; c'est encore un problème de dynamique, de sorte que la difficulté principale subsiste ; mais c'est le plus simple de tous les problèmes de dynamique ; d'abord il n'y a que deux degrés de liberté, et puis si l'on prend une surface sans point singulier, on n'a rien de comparable avec la difficulté que l'on rencontre dans les problèmes de dynamique aux points où la vitesse est nulle ; dans le problème des lignes géodésiques, en effet, la vitesse est constante et peut donc être regardée comme une des données de la question.*

2.1. **Existence of closed geodesics.** In this section, M is a closed connected smooth manifold, endowed with a Riemannian metric g. Two closed cuves on M are *freely homotopic* if they are homotopic. (The word is introduced to distinguish from pointed homotopy classes, that form the fundamental group of M.) Free homotopy defines an equivalence relation on the free loop space ΛM. The equivalence classes are the path components of ΛM. They are called *free homotopy classes*, and we denote the set of free homotopy classes by $\mathcal{F}(M)$. The class of contractible curves contains the point curves. In this section we give the most elementary proof known to us of the following theorem.

Theorem 2.1. *Let M be a closed manifold and let $\alpha \in \mathcal{F}(M)$ be a free homotopy class of closed curves.*

(i) *If α is non-trivial, then there exists a closed geodesic in class α.*

(ii) *If $\mathcal{F}(M)$ contains only the class of contractible curves, then there exists a contractible closed geodesic on M.*

In particular, every closed Riemannian manifold carries a closed geodesic.

Proof. The distance $d(p, q)$ between two points in M is defined as the infimum of the length of a smooth curve from p to q. We endow the loop space ΛM with the C^0-topology, defined by the metric

$$d(c_1, c_2) = \max_{t\in[0,1]} d(c_1(t), c_2(t)).$$

Recall that there is a number $\rho_0 > 0$ with the property that any two points $p, q \in M$ with $d(p,q) \le \rho_0$ can be connected by a unique geodesic of shortest length. Moreover, this geodesic depends continuously on p and q. (One can take as ρ_0 any number smaller than the injectivity radius of (M,g).

Lemma 2.2. *Let $c_0, c_1 \colon S^1 \to M$ be two curves with $d(c_0,c_1) \le \rho_0$. Then c_0 and c_1 are homotopic.*

Proof. For each $t \in S^1$ let $\gamma_t(s) \colon [0,1] \to M$ be the unique geodesic from $c_0(t)$ to $c_1(t)$, parametrized proportional to arc-length. Since γ_t depends continuously on its endpoints, it depends continuously on t. Hence

$$\Gamma(t,s) := \gamma_t(s), \quad t \in S^1,\ s \in [0,1],$$

is continuous and yields the desired homotopy from c_0 to c_1, see Figure 6. □

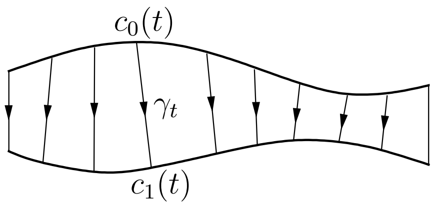

FIGURE 6. The homotopy Γ.

Proof of (i). Let ℓ_α be the infimum of the lengths of piecewise differentiable curves belonging to α. Since α is non-trivial, $\ell_\alpha > 0$. Let $c_1, c_2, \dots$ be a sequence of piecewise differentiable curves in α whose lengths $\ell(c_j)$ are monotone decreasing to ℓ_α. Choose $N \in \mathbb{N}$ such that $\ell(c_1)/N \le \rho_0$. Consider the partition $0 = t_0 < t_1 < \dots < t_N = t_0 = 1$ of S^1 with $t_{i+1} - t_i = \frac{1}{N}$ for all i. We can assume that the curves c_j are parametrized proportional to arc-length. Then

$$d(c_j(t_i), c_j(t_{i+1})) \le \ell(c_j|_{[t_i,t_{i+1}]}) \le \ell(c_j)/N \le \rho_0 \quad \text{for all } j \text{ and } i. \tag{1}$$

Since M is compact, its N-fold product M^N is also compact. The sequence $(c_j(t_0), \dots, c_j(t_{N-1}))$ in M^N therefore has a convergent subsequence, that we still denote $(c_j(t_0), \dots, c_j(t_{N-1}))$. Let $(p_0, \dots, p_{N-1})$ be its limit. Then $d(p_i, p_{i+1}) = \lim_{j\to\infty} d(c_j(t_i), c_j(t_{i+1})) \le \rho_0$ in view of (1). Let γ be the broken geodesic with vertices p_i. Using again (1) we find

$$\begin{aligned}
\ell(\gamma) = \sum_{i=0}^{N-1} d(p_i,p_{i+1}) &= \sum_{i=0}^{N-1} \lim_{j\to\infty} d(c_j(t_i), c_j(t_{i+1})) \\
&= \lim_{j\to\infty} \sum_{i=0}^{N-1} d(c_j(t_i), c_j(t_{i+1})) \le \lim_{j\to\infty} \ell(c_j) = \ell_\alpha
\end{aligned}$$

and hence $\ell(\gamma) = \ell_\alpha$.

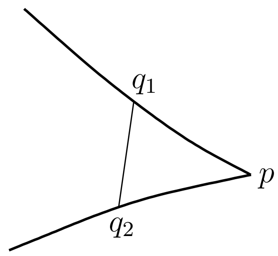

FIGURE 7. Constructing a shorter curve.

It remains to show that γ is smooth at each vertex p_i. Assume the contrary, namely that γ is not smooth at some point p. Choose points q_1, q_2 on γ as in Figure 7 such that $d(q_1, p)$, $d(p, q_2)$ and $d(q_1, q_2)$ are at most ρ_0 and such that the curve $q_1 p\, q_2$ is homotopic to the unique geodesic $q_1 q_2$ of length $\leq \rho_0$, relative end points. Replacing the part $q_1 p\, q_2$ of γ by $q_1 q_2$ we obtain a closed piecewise smooth curve $\tilde{\gamma}$ with $\ell(\tilde{\gamma}) < \ell(\gamma) = \ell_\alpha$ that also lies in class α. This contradicts the definition of ℓ_α.

Proof of (ii). We first assume that M is a 2-sphere. Recall that we identify the circle S^1 with $\mathbb{R}/\mathbb{Z}$. Denote the standard sphere in $\mathbb{R}^3$ of radius 1 by $\mathbb{S}^2$. Choose a smooth map $f\colon \mathbb{S}^2 \to S^2 =: M$ of degree one: $[f(\mathbb{S}^2)] = [M] = 1 \in \mathbb{Z} \cong \pi_2(M)$. The map f induces a "tapestry" on M, namely the family of curves $\{c_s\colon S^1 \to M\}$, $c \in [-1, 1]$. Now let's pull along all these curves! Let $v := \max_{-1 \leq s \leq 1} \|\dot{c}_s(t)\|_\infty$ be the maximal velocity of a curve c_s. Denote by $\mathcal{L}^v$ the space of piecewise smooth curves $c\colon S^1 \to M$ for which $\|\dot{c}(t)\| \leq v$ at smooth points of c, and let $\mathcal{B}^v$ be the subset of $\mathcal{L}^v$ consisting of broken geodesics. The length of $c \in \mathcal{L}^v$ is, of course, the sum of the lengths of its smooth pieces.

The main tool of the proof is a curve shortening procedure:

Proposition 2.3. *There exists a map $P\colon \mathcal{L}^v \to \mathcal{B}^v$ with the following properties.*

(i) $\ell(P(c)) \leq \ell(c)$ *with equality if and only if c is either a geodesic or a point curve;*
(ii) *for every c there exists a subsequence $n_k \subset \mathbb{N}$ such that $P^{n_k}(c)$ converges uniformly to either a geodesic or a point curve;*
(iii) *the map $P\colon \mathcal{L}^v \to \mathcal{B}^v$ is homotopic to the identity of $\mathcal{L}^v$;*
(iv) *the map P is continuous in the C^0-topology, and the functions*

$$f_n\colon [0,1] \to \mathbb{R}, \quad s \mapsto \ell(P^n(c_s))$$

are continuous.

Proof. Recall that any two points $p, q \in M$ with $d(p,q) \leq \rho_0$ can be connected by a unique geodesic of shortest length. Moreover, this geodesic depends continuously on p and q. Choose N so large that $L/N \leq \rho_0$. Consider the partitions $0 = t_0 < t_1 < \cdots < t_N = t_0 = 1$ and $\frac{1}{2N} = \tau_1 < \cdots < \tau_{2N+1} = \frac{2N+1}{2N}$ of S^1 with $t_i - t_{i-1} = \frac{1}{N}$ and $\tau_i - \tau_{i-1} = \frac{1}{N}$ for all i, and $t_0 = t_N$, $\tau_1 = \tau_{2N+1}$. For $c \in \mathcal{L}^v$ we can then define $P_1(c)$ (resp. $P_2(c)$) to be the unique broken geodesic with vertices $c(t_i)$ (resp. $c(\tau_i)$). Set $P := P_2 \circ P_1$.

Exercise 2.4. Verify that P meets properties (i)–(iv).

By (ii) we find for every $s \in [-1,1]$ a subsequence $n_k(s)$ such that $P^{n_k(s)}(c_s)$ converges uniformly to a closed geodesic or to a point curve. Arguing by contradiction, assume now that for each s this limit curve is a point. By (iv), the functions $f_n(s) = \ell(P^n(c_s))$ are continuous, and by (i) and by our assumption, for each s the sequence $f_n(s)$ is monotone and tends to 0. In other words, the sequence f_n converges monotone and pointwise to the continuous function $f(t) \equiv 0$. Dini's theorem thus implies that $f_n \to 0$ uniformly on $[-1,1]$. Hence there exists n_0 such that $\ell(P^{n_0}(c_s)) \leq \rho_0$ for all $s \in [-1,1]$.

Exercise 2.5. Conclude that $f(S^2)$ is contractible in M.

This contradiction shows that there exists s_0 such that a subsequence of $P^n(c_{s_0})$ tends to a closed geodesic. □

Remark 2.6. 1. The proof of assertion (i) of Theorem 2.1 yields a geodesic in class α that minimizes the length of all rectifiable closed curves in class α. The curve shortening procedure used in the proof of (ii), applied to a curve in a non-trivial class α, also yields the existence of a closed geodesic in class α. This geodesic may, however, not minimize the length of curves in α.

2. Most authors prove Theorem 2.1 (ii) by considering the energy functional $\mathcal{E}(c) = \frac{1}{2}\int_0^1 \|\dot{c}(t)\|^2\,dt$ instead of the length $\mathcal{L}(c)$. The energy functional has the advantage that for generic Riemannian metrics (for which the closed geodesics are isolated) the components of the critical set of $\mathcal{E}$ (namely the closed geodesics) are compact (namely circles), see Proposition **??** wie Dietmar. This is relevant for doing Morse theory on the free loop space. On the other hand, the length of a closed geodesic does not change under reparametrization, and so a component of the critical set of $\mathcal{L}$ is at least as large as the group of orientation preserving diffeomorphisms of the circle, which is not compact. For the proof of Theorem 2.1 (ii) this disadvantage of the length functional is irrelevant, and we found it more natural and also slightly more elementary to work with the length.

3. Assertion (i) of Theorem 2.1 goes back to Cartan and Hadamard. Assertion (ii) was proven by G. Birkhoff in 1917 for S^2 and in 1927 [4] for S^n. The general case was proven by Lusternik and Fet in 1947 [8]. Hurewicz's theorem used in the proof of Lemma **??** was not available in 1927.

Exercise 2.7. (i) Consider the circle $S^1 \subset \mathbb{C}$, parametrized by $\gamma(t) = e^{2\pi i t}$. Fix $k \geq 1$, and choose the two partitions $t_j = j/k$, $\tau_j = (j - 1/2)/k$. Show that $P^n(\gamma)$ converges to the origin.

(ii) Decompose $S^2 \subset \mathbb{R}^3$ into loops $c_s = S^2 \cap \{x_1 = s\}$, $s \in [0,1]$. Parametrize each loop on $[0,1]$ proportional to arc-length. Describe the family $P^n(c_s)$ for n large.

2.2. **Counting closed geodesics.** It is conjectured, and known for most closed manifolds, that for every Riemannian metric g on M there exists infinitely many closed geodesics.

Once this is known, one wants to count the closed geodesics finer: below length. In view of Theorem 2.1 (i) we first study the set $\mathcal{F}(M)$ of free homotopy classes.

Let $\sim$ be the equivalence relation on $\pi_1(M)$ given by conjugacy: $\alpha \sim \beta$ if there exists $\gamma \in \pi_1(M)$ such that $\alpha = \gamma\beta\gamma^{-1}$.

Exercise 2.8. *The set of free homotopy classes $\mathcal{F}(M)$ is in bijection with $\pi_1(M)/\sim$.*

2.3. **Geodesic and horocycle dynamics on the hyperbolic plane and its compact quotients.**

2.3.1. *Geodesics and horocycles on the upper half plane.* Endow the upper half plane

$$\mathbf{H} := \{z = x + iy \in \mathbb{C} \mid y > 0\}$$

with the Riemannian metric

$$ds^2 = \frac{1}{y^2}\left(dx^2 + dy^2\right). \tag{2}$$

This Riemannian manifold is, up to isometry, the unique simply connected surface of constant curvature -1. The group

$$\mathrm{PSL}(2,\mathbb{R}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,b,c,d \in \mathbb{R};\ ad - bc = 1 \right\} / \{\pm 1\}$$

acts biholomorphically on $\mathbf{H}$ via the mappings

$$z \mapsto \frac{az+b}{cz+d}$$

and leaves the metric (2) invariant. In fact, $\mathrm{PSL}(2,\mathbb{R})$ is the full group of orientation preserving isometries of $\mathbf{H}$. Typical examples are the inversion $z \mapsto -1/z$, the translations $z \mapsto z + b$ ($b \in \mathbb{R}$), and the dilations $z \mapsto az$ ($a > 0$).

The geodesics on $\mathbf{H}$ are precisely the vertical lines and the half-circles with center on the real axis. The horizontal lines $\mathbb{R}+iy = \{r+is \mid r \in \mathbb{R}\}$ are called *horocycles* centered at ∞; the intersection of circles tangent to $\mathbb{R}$ at $x \in \mathbb{R}$ with $\mathbf{H}$ are called *horocycles* centered at x, see Figure 8.

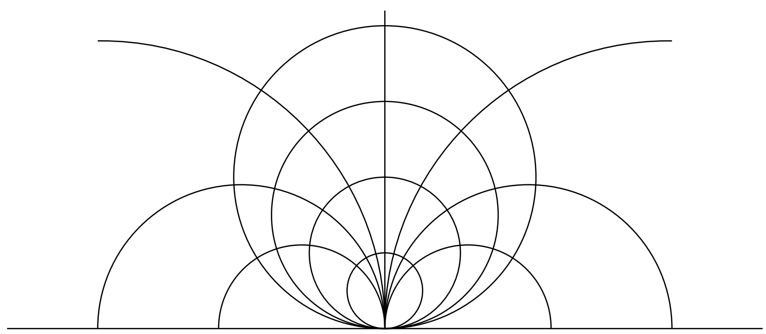

FIGURE 8. Geodesics and horocycles on $\mathbb{H}$.

Lemma 2.9. *For every horocycle H there exists an orientation preserving isometry T of* **H** *such that $T(\mathbb{R}+i)=H$.*

Proof. For horocycles $H=\mathbb{R}+is$ take a dilation $Tz=sz$. The inversion $Tz=-1/z$ maps $\mathbb{R}+i$ to the horocycle centered at $(0,0$ and passing through i: For all $r\in\mathbb{R}$ we have

$$\left|T(r+i)-\frac{i}{2}\right| = \left|\frac{1}{r+i}+\frac{i}{2}\right| = \left|\frac{2+i(r+i)}{2(r+i)}\right| = \left|\frac{i(r-i)}{2(r+i)}\right| = \frac{1}{2}.$$

The horocycle centered at x of (Euclidean) radius a is obtained from this horocycle by first applying the dilation by a and then the translation by x. □

ev hier geodesic curvature ...

2.4. **Compact quotients.** Now let Γ be a discrete subgroup of $\mathrm{PSL}(2,\mathbb{R})$ that acts on **H** without fixed points. Then the quotient $M_\Gamma := \Gamma\setminus\mathbf{H}$ is a smooth manifold, and the Riemannian metric (2) descends to M_Γ. Then the projection map is a local isometry. Therefore, the geodesics on M_Γ are the projections of geodesics on **H**, the curves of geodesic curvature $+1$ are the projections of the horocycles on **H** (which are called again horocycles), and the Gauss curvature of M_Γ is -1. From now on, we also assume that M_Γ is compact. By the Gauss–Bonnet theorem, we then must have

$$k\,\mathrm{Vol}\,M_\Gamma \;=\; -\,\mathrm{Vol}\,M_\Gamma \;=\; 2\pi\,\chi(M_\Gamma)$$

where $\chi(M_\Gamma)$ is the Euler characteristic of M_Γ. Since the Euler characteristic of a surface of genus g is $2-2g$, we see that we must have $g\geq 2$. That for $g\geq 2$ such subgroups Γ exist is not obvious. For a few explicit examples see [10, p. 215], [5], [**?**]. In turns out that such examples are abundant: The set of closed 2-dimensional Riemannian manifolds modulo isometries isotopic to the identity (the so-called Teichmüller space) is homeomorphic to $\mathbb{R}^{6(g-1)}$ (see [2, B.4], [5, Ch. 6].

Lemma 2.10. *No horocycle in M_Γ is closed.*

Proof. Let H be a horocycle in M_Γ, parametrized by arc-length, and let $\widetilde{H}$ be its lift to **H**. Arguing by contradiction we assume that H is closed. Since the horocycle $\widetilde{H}$ is not a closed curve, H is a non-contractible closed curve in M_Γ. Therefore, there exists a Deck transformation $\varphi\in\pi_1(M_\Gamma)\cong\Gamma\subset\mathrm{PSL}(2;\mathbb{R})$ with $\varphi\neq id$ and $\varphi(\widetilde{H})=\widetilde{H}$. Recall from Theorem 2.1 that the free homotopy class corresponding to $\varphi\in\Gamma\cong\pi_1(M_\Gamma)$ contains a closed geodesic.[1] From this it follows easily that "φ has an axis", i.e., there exists a geodesic c on **H** such that $\varphi(c)=c$, see for instance [6, Chap. 12, Prop. 2.6]. By Lemma (2.9) there exists an isometry $T\in\mathrm{PSL}(2;\mathbb{R})$ such that $T(H)=\mathbb{R}+i$. Then the isometry $\psi:=T\circ\varphi\circ T^{-1}\in\mathrm{PSL}(2;\mathbb{R})$ leaves invariant the geodesic $T\circ c$ and the

[1] Here we use that M_Γ is closed. The generator of the fundamental group of the pseudosphere $\Gamma\setminus\mathbf{H}$, where Γ is generated by the translation $T(z)=z+1$, cannot be represented by a closed geodesic.

horocycle $\mathbb{R}+i$. Since ψ is an isometry, it translates $\mathbb{R}+i$ by a constant, say τ. If $\psi(z)=\frac{az+b}{cz+d}$, we thus have

$$\psi(r+i) = \frac{a(r+i)+b}{c(r+i)+d} = r+i+\tau \quad \text{for all } r\in\mathbb{R}.$$

It readily follows that ψ is a translation of $\mathbf{H}$, $\psi(z)=z+b=z+\tau$. Since $\varphi\neq id$, also $\psi\neq id$. However, a non-trivial translation leaves invariant no geodesic of $\mathbf{H}$. □

Remark 2.11. In the above proof we followed [9, p. 137]. A seminal result of Hedlund [] is much stronger: The horocycle flow on M_Γ is minimal, i.e., each orbit of the horocycle flow is dense in T_1M_Γ. In particular, each horocycle is dense in M_Γ (and hence not closed).

As a introduction to hyperbolic geometry we recommend [2]. All we needed on hyperbolic geometry can be found in [10, Ch. 5.4]. The traditional way is to start from the Riemannian metric, and then to compute the geodesics and isometries. A more intrinsic approach starts from geodesics, and then computes the isometries and the metric, see [14, ???].

3. Solutions of exercises

Exercise 1.2. We give two proofs, the first based on the Jordan curve theorem, the second on the notion of intersection degree.

First proof. Assume that γ_1,γ_2 do no intersect. Since their images are compact, their distance is then positive. We can therefore approximate γ_1,γ_2 by piecewise linear curves $\tilde\gamma_1,\tilde\gamma_2$ that are still disjoint and have finitely many and transverse self-intersections. After removing finitely many closed arcs from $\tilde\gamma_1$, we can assume that $\tilde\gamma_1$ is embedded. We can also assume that $\tilde\gamma_1$ avoids the points $(0,1)$ and $(1,0)$. (and piecewise linear ? for Jordan?) By the Jordan curve theorem, the image of $\tilde\gamma_1$ divides $Q=[0,1]^2$ into two regions R_1 containing $(0,1)$ and R_2 containing $(1,0)$. By the Tietze extension theorem we find a continuous function $f\colon Q\to[-1,1]$ with $f^{-1}(0)=\tilde\gamma_1$ and such that f is negative on $R_1\setminus\tilde\gamma_1$ and positive on $R_2\setminus\tilde\gamma_1$. Now consider the continuous function $g\colon[0,1]\to[-1,1]$, $g(t)=f\circ\tilde\gamma_2(t)$. Since $\tilde\gamma_2$ starts in $(0,1)\in R_1\setminus\tilde\gamma_1$ and ends in $(1,0)\in R_2\setminus\tilde\gamma_1$, the intermediate value theorem implies that there exists t^* with $g(t^*)=0$, i.e., $\tilde\gamma_2(t^*)\subset\tilde\gamma_1$. Hence the curves $\tilde\gamma_1,\tilde\gamma_2$ intersect.

Second proof. Given a continuous curve $\gamma_1\colon[0,1]\to Q$ from $(0,0)$ to $(1,1)$ and a continuous curve $\gamma_2\colon[0,1]\to Q$ from $(0,1)$ to $(1,0)$, the $\mathbb{Z}_2$-intersection degree $\deg(\gamma_1,\gamma_2)\in\{0,1\}$ is defined as follows. If γ_1 and γ_2 are smooth and intersect transversally, then $\deg(\gamma_1,\gamma_2)$ is the number modulo 2 of intersections of γ_1 and γ_2. In general, $\deg(\gamma_1,\gamma_2):=\deg(\tilde\gamma_1,\tilde\gamma_2)$, where $\tilde\gamma_1,\tilde\gamma_2$ are smooth approximations of γ_1,γ_2 that intersect transversally. This number is indeed well-defined, and it is invariant under homotopies of γ_1,γ_2 relative endpoints.

Let now γ_1,γ_2 be the two curves in our problem. Since γ_1 is homotopic to the diagonal $\tilde\gamma_1$ from $(0,0)$ to $(1,1)$ and γ_2 is homotopic to the anti-diagonal $\tilde\gamma_2$ from $(0,1)$ to $(1,0)$, $\deg(\gamma_1,\gamma_2)=\deg(\tilde\gamma_1,\tilde\gamma_2)=1$. Hence γ_1,γ_2 intersect.

Exercise 1.4. We propose two solutions, a more topological and a more algebraic one.

Solution 1. The space $\mathbb{R}\mathrm{P}^3$ is defined as the quotient $S^3/\{x=-x\}$. By restricting the antipodal action on S^3 to a closed hemisphere, we see that $\mathbb{R}\mathrm{P}^3$ is obtained from a closed 3-ball by identifying antipodal boundary points.

Consider the closed ball $B^3(\pi) \subset \mathbb{R}^3$ of radius π. Define the map $\varphi\colon B^3(\pi) \to \mathbf{SO}(3)$ by sending a nonzero vector x to the rotation through angle $|x|$ about the axis formed by the line through the origin in the direction of x. Here, we use the "right-hand-rule". By continuity, φ then sends 0 to the identity. Antipodal points of $\partial B^3(\pi)$ are sent to the same rotation through angle π, so φ induces a map $\overline{\varphi}\colon\ \mathbb{R}\mathrm{P}^3 \to \mathbf{SO}(3)$. The map $\overline{\varphi}$ is injective, because the axis of a nontrivial rotation is uniquely determined as its fixed point set, and $\overline{\varphi}$ is surjective, because every non-identity element of $\mathbf{SO}(3)$ has 1 as a simple eigenvalue, and hence is a rotation about some axis. It follows that $\overline{\varphi}$ is a homeomorphism.

It is not easy to see whether $\overline{\varphi}$ is a diffeomorphism. That the spaces $\mathbb{R}\mathrm{P}^3$ and $\mathbf{SO}(3)$ are diffeomorphic follows from a theorem of Bing and Moise according to which every topological 3-manifolds has a unique smooth structure. In the following solution, the smoothness of the map is clear.

Solution 2. Identify $S^3 \subset \mathbb{R}^4 = \mathbb{H}$ with the set of quaternions $a+bi+cj+dk$ of length 1. Multiplication of quaternions satisfies $|xy| = |x|\,|y|$. View $\mathbb{R}^3$ as the set of "pure imaginary" quaternions $\{bi+cj+dk\}$.

(a) For $x \in S^3$ the map $y \mapsto xyx^{-1}$ defines an orthogonal map $f_x\colon \mathbb{R}^3 \to \mathbb{R}^3$ with determinant $+1$. We thus obtain a map $f\colon S^3 \to \mathbf{SO}(3)$, $x \mapsto f_x$.

(b) The map $f\colon S^3 \to \mathbf{SO}(3)$ is smooth, surjective, and $f(x) = f(x')$ if and only if $x' = \pm x$.

(c) The map $f\colon S^3 \to \mathbf{SO}(3)$ is a 2-fold covering map.

(d) f induces a diffeomorphism $\mathbb{R}\mathrm{P}^3 \to \mathbf{SO}(3)$.

Exercise 1.5. Label the three arms of the linkage by 1, 2, 3. The usual orientation of the plane $\mathbb{R}^2$ induces an orientation of the hexaxon H. The configuration space Q is obtained from gluing the eight hexagons

$$H_{LLL},\ H_{LLR},\ H_{LRL},\ H_{RLL},\ H_{LRR},\ H_{RLR},\ H_{RRL},\ H_{RRR}$$

where, for instance, H_{LLR} is the hexagon whose interior points correspond to configurations with the first and second elbow turned to the left and the third elbow turned to the right. Label each edge of such a hexagon by one of $1^\pm$, $2^\pm$, $3^\pm$, where 1^- (resp. 1^+) is the edge lying on the smaller (resp. larger) circle centred at 1, etc.

figure

The eight hexagons are glued as follows: Two hexagons have a common edge if and only if they differ in exactly one of the three letters from the set $\{L, R\}$, say in position j, and in this case the two edges j^- are glued and the two edges j^+ are glued. E.g., H_{LLR} and H_{LRR} are glued along the two edges 2^- and the two edges 2^+.

Note that the gluing along edges is done abstractly. E.g. H_{LLR} is glued to H_{LRR} by putting $H_{LLR} \subset \mathbb{R}^2$ on top of $H_{LRR} \subset \mathbb{R}^2$ and identifying the edges 2^- and the edges 2^+. On the other hand, a smooth curve in Q crossing an edge transversely corresponds to a curve in $H \subset \mathbb{R}^2$ that is reflected at the corresponding edge. It is now clear how to define an orientation of Q: Orient a hexagon $H \subset Q$ like H if none or two of its labels are L, and orient $H \subset Q$ opposite to H if one or three of its labels differ from L. For instance, $H_{LLR} \subset Q$ is oriented like $H_{LLR} \subset \mathbb{R}^2$, while $H_{LRR} \subset Q$ is oriented opposite to $H_{LRR} \subset \mathbb{R}^2$. This defines an orientation of Q because any sequence of moves from a given word in $\{L, R\}$ with three letters to another given such word (where each move changes exactly one letter) either has even or odd length. Equivalently, any closed curve in Q (that we may assume to avoid vertices and to intersect edges transversally) crosses an even number of edges, and hence the orientation along closed curves is preserved.

Exercise 2.8. Fix a point $p \in M$. Consider the maps

$$\mathcal{F}(M) \xrightarrow{\varphi} \pi_1(M,p)/\sim \qquad \text{and} \qquad \mathcal{F}(M) \xleftarrow{\psi} \pi_1(M,p)/\sim$$

defined as follows. For $[\alpha] \in \mathcal{F}(M)$ choose a closed curve $\tilde{\alpha}$ based at p that is freely homotopic to α, and set $\varphi([\alpha]) = [[\tilde{\alpha}]]$, where $[[\tilde{\alpha}]]$ is the projection of $[\tilde{\alpha}] \in \pi_1(M,p)$ to $\pi_1(M,p)/\sim$. Conversely, given an element $[[\beta]] \in \pi_1(M)/\sim$ with $[\beta] \in \pi_1(M,p)$, define $\psi([[\beta]]) = [\beta] \in \mathcal{F}(M)$. We only need to show that φ and ψ are well-defined, since then $\psi \circ \varphi = \mathrm{id}_{\mathcal{F}(M)}$ and $\varphi \circ \psi = \mathrm{id}_{\pi_1(M,p)/\sim}$.

φ is well-defined: Let $h(t,s)\colon [0,1] \times [0,1] \to M$ be a free homotopy between closed curves $\alpha_1, \alpha_2\colon [0,1] \to M$ based at p. Then

$$h(t,0) = \alpha_1(t),\ h(t,1) = \alpha_2(t) \quad \text{and} \quad \gamma(s) := h(0,s) = h(1,s) \ \text{ for all } \ t \in [0,1].$$

In particular, $[\gamma] \in \pi_1(M,p)$.

figure

By restricting h to the family of paths indicated in Figure **??** we see that α_1 is homotopic to $\gamma^{-1} \circ \alpha_2 \circ \gamma$ relative p, that is, $[\alpha_1] = [\gamma]^{-1}[\alpha_2][\gamma]$ in $\pi_1(M,p)$.

ψ is well-defined: Assume that β_1, β_2 and γ are closed curves based at p such that β_2 is homotopic to $\gamma^{-1} \circ \beta_1 \circ \gamma$ relative p. Moving the base point 0 of $[0,1]/\{0,1\}$ by 1/3 we see that $\gamma^{-1} \circ \beta_1 \circ \gamma$ is freely homotopic to $\gamma \circ \gamma^{-1} \circ \beta_1$, which is homotopic to β_1 relative p. Hence β_2 is freely homotopic to β_1, i.e., $[\beta_1] = [\beta_2] \in \mathcal{F}(M)$.

References

[1] Arnol'd, V. I. *Mathematical methods of classical mechanics.* Second edition. Graduate Texts in Mathematics **60**. Springer-Verlag, New York, 1989.

[2] R. Benedetti and C. Petronio. *Lectures on hyperbolic geometry.* Universitext. Springer-Verlag, Berlin, 1992.

[3] G. Birkhoff. Dynamical systems with two degrees of freedom. *Trans. Amer. Math. Soc.* **18** (1917) 199–300.

[4] G. Birkhoff. *Dynamical systems.* With an addendum by Jürgen Moser. AMS Colloquium Publications, Vol. IX. AMS, Providence, 1966. Reprinting of the original edition [New York, 1927].

[5] P. Buser. *Geometry and spectra of compact Riemann surfaces.* Modern Birkhäuser Classics. Birkhäuser, Boston, 2010.

[6] M. P. do Carmo. *Riemannian geometry.* Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992.

[7] M. Farber. *Invitation to topological robotics.* Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008.

[8] L. A. Lyusternik and A. I. Fet. Variational problems on closed manifolds. *Doklady Akad. Nauk SSSR* **81** (1951) 17–18.

[9] V. Ginzburg. On closed trajectories of a charge in a magnetic field. An application of symplectic geometry. Contact and symplectic geometry (Cambridge, 1994), 131–148, *Publ. Newton Inst.* **8**, Cambridge Univ. Press, Cambridge, 1996.

[10] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems.* With a supplementary chapter by Katok and Leonardo Mendoza. *Encyclopedia of Mathematics and its Applications* **54**. Cambridge University Press, Cambridge, 1995.

[11] W. Massey. *Algebraic topology: an introduction.* Graduate Texts in Mathematics **56**. Springer–Verlag, New York-Heidelberg, 1977.

[12] H. Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta Mathematica* **13** (1890) 1–270.

[13] H. Poincaré. Sur les lignes géodésiques des surfaces convexes. *Trans. Amer. Math. Soc.* **6** (1905) 237–274.

[14] W. Thurston. *Three-dimensional geometry and topology. Vol. 1.* Princeton Mathematical Series **35**. Princeton University Press, Princeton, NJ, 1997.

Felix Schlenk, Institut de Mathématiques, Université de Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

E-mail address: `schlenk@unine.ch`