Math 53H: The Lie derivative

Yakov Eliashberg

May 19, 2011

1 Lie derivative of a differential form

Let A be a smooth vector field defined on a domain $U \subset \mathbb{R}^n$ (more generally we can assume that U is any n-dimensional manifold). Given a function $f : U \to \mathbb{R}$ we can define the *directional derivative* $L_A f$ of f along A:

$$L_A f = \lim_{s \to 0} \frac{f(x + tX) - f(x)}{t}. \tag{1.1}$$

The directional derivative has many other notation: $D_A(f), \frac{\partial f}{\partial A}, df(A), \dots$.

Let us denote by $A^t : U \to U$, $t \in \mathbb{R}$, the phase flow of the vector field A. [1] Let us observe that the directional derivative can be also defined by the formula

$$L_A f = \frac{d}{ds} f \circ A^s \Big|_{s=0}. \tag{1.2}$$

It turns out that formula (1.2) can be generalized to define an analog of directional derivatives for differential forms and vector fields, which is the *Lie derivative.*

Let ω be a differential k-form. We define the Lie derivative $L_A\omega$ of ω along A as

$$L_A \omega = \frac{d}{ds} (A^s)^* \omega \Big|_{s=0}. \tag{1.3}$$

[1] Note that that the phase flow is not necessarily globally defined, and may mot be defined for all t. However all our considerations in this section will be local, and according to the existence and uniqueness theorem for ODE, the flow is always locally defined. Hence, to simplify the notation, we will not be making this distinction between the local and global.

Note that for if ω is a 0-form, i.e. a function f, then $(A^s)^* f = f \circ A^s$, and hence, in this case definitions (1.2) and (1.3) coincide, and hence for functions the Lie derivative is the same as the directional derivative.

Proposition 1.1. *The following identities hold*

1. $L_A(\omega_1 \wedge \omega_2 = (L_A\omega_1) \wedge \omega_2 + \omega_1 \wedge L_A\omega_2$.

2. $L_A(d\omega) = d(L_A\omega)$.

Proof.

1. $$L_A(\omega_1 \wedge \omega_2) = \frac{d}{ds}(A^s)^*(\omega_1 \wedge \omega_2)\Big|_{s=0} = \frac{d}{ds}\big((A^s)^*\omega_1 \wedge (A^s)^*\omega_2\big)\Big|_{s=0}$$
$$= \frac{d}{ds}\big((A^s)^*\omega_1\big)\Big|_{s=0} \wedge \omega_2 + \omega_1 \wedge \frac{d}{ds}\big((A^s)^*\omega_2\big)\Big|_{s=0} = (L_A\omega_1) \wedge \omega_2 + \omega_1 \wedge L_A\omega_2.$$

2. $$L_A(d\omega) = \frac{d}{ds}\big((A^s)^* d\omega\big)\Big|_{s=0} = \frac{d}{ds}\big(d(A^s)^*\omega\big)\Big|_{s=0} = d\Big(\frac{d}{ds}(A^s)^*\omega\Big|_{s=0}\Big) = L_A(d\omega).$$

■

The following formula of Eli Cartan provides an effective way for computing Lie derivative of differential form.

Theorem 1.2. *Let A be a vector field and ω a differential k-form. Then*

$$L_A\omega = d(A \lrcorner\, \omega) + A \lrcorner\, d\omega. \tag{1.4}$$

Proof. Suppose first that $\omega = f$ is a 0-form. Then $L_A f = df(A) = A \lrcorner df$, which is equivalent to formula (1.4), because in this case the first term in the formula is equal to 0. Then, using Proposition 1.12) we get

$$L_A df = dL_A f = d(df(A)) = d(A \lrcorner\, df),$$

which is again equivalent to (1.4) because in this case $ddf = 0$. Next we note that if the formula (1.1) holds for ω_1 and ω_2 then it holds also for $\omega_1 \wedge \omega_2$. Indeed, denothe the degree of forms ω_1, ω_2 by d_1, d_2. Then we have

$$(\star)\ \ L_A(\omega_1 \wedge \omega_2) = (L_A\omega_1) \wedge \omega_2 + \omega_1 \wedge L_A\omega_2$$

$$= (A \lrcorner d\omega_1 + d(A \lrcorner \omega_1)) \wedge \omega_2 + \omega_1 \wedge (A \lrcorner d\omega_2 + d(A \lrcorner \omega_2))$$

$$= (A \lrcorner d\omega_1) \wedge \omega_2 + \omega_1 \wedge (A \lrcorner d\omega_2) + d(A \lrcorner \omega_1) \wedge \omega_2 + \omega_1 \wedge d(A \lrcorner \omega_2)$$

On the other hand

$$(\star\star)\ A \lrcorner d(\omega_1 \wedge \omega_2) + d(A \lrcorner (\omega_1 \wedge \omega_2))$$

$$= A \lrcorner (d\omega_1 \wedge \omega_2 + (-1)^{d_1}\omega_1 \wedge d\omega_2) + d((A \lrcorner \omega_1) \wedge \omega_2 + (-1)^{d_1}\omega_1 \wedge (A \lrcorner \omega_2))$$

$$= (A \lrcorner d\omega_1) \wedge \omega_2 + (-1)^{d_1+1} d\omega_1 \wedge (A \lrcorner \omega_2) + (-1)^{d_1}(A \lrcorner \omega_1) \wedge d\omega_2 + \omega_1 \wedge (A \lrcorner d\omega_2)$$

$$+ d(A \lrcorner \omega_1) \wedge \omega_2 + (-1)^{d_1+1} A \lrcorner \omega_1 \wedge d\omega_2 + (-1)^{d_1} d\omega_1 \wedge (A \lrcorner \omega_2) + \omega_1 \wedge (d(A \lrcorner \omega_2))$$

$$= (A \lrcorner d\omega_1) \wedge \omega_2 + \omega_1 \wedge (A \lrcorner d\omega_2) + d(A \lrcorner \omega_1) \wedge \omega_2 + \omega_1 \wedge d(A \lrcorner \omega_2)\,.$$

Comparing the computation in $(\star)$ and $(\star\star)$ we conclude that

$$L_A(\omega_1 \wedge \omega_2) = A \lrcorner d(\omega_1 \wedge \omega_2) + d(A \lrcorner (\omega_1 \wedge \omega_2)).$$

By induction we can prove a similar formulas for an exterior product of any number of forms.

Finally we observe that any differential k-form ω can be written in coordinates as $\sum\limits_{1 \le i_1 < \cdots < i_k \le n} f_{i_1 \ldots i_k}(x) dx_{i_1} \wedge \cdots \wedge dx_{i_k}$, i.e. ω is a sum of exterior products of functions (0-forms) and exact 1-forms, and hence Cartan's formula follows. ■

Proposition 1.3. *We have*

$$L_A\omega = 0 \iff (A^s)^*\,\omega = \omega \ \ \textit{for all}\ \ s \in \mathbb{R}.$$

Proof. If $(A^s)^*\,\omega \equiv \omega$ then $L_A\omega = \frac{d}{ds}(A^s)^*\omega\Big|_{s=0} = 0$. To prove the converse we note that

$$\frac{d}{ds}(A^s)^*\omega\Big|_{s=s_0} = \lim_{t\to 0} \frac{(A^{s_0+t})^*\omega - (A^{s_0})^*\omega}{t} = (A^{s_0})^* \left(\lim_{t\to 0} \frac{(A^t)^*\omega - \omega}{t} \right)$$
$$= (A^{s_0})^*\,(L_A\omega)\,,$$

and hence if $L_a\omega = 0$ then $(A^s)^*\,\omega = \omega$. ■

2 Action of diffeomorphisms on vector fields

Let $U \subset \mathbb{R}^n$ be a domain in $\mathbb{R}^n$ and $f : U \to V$ be a diffeomorphism of U onto another domain $V \subset \mathbb{R}^n$. We define the *push-forward* operator $f_* : \mathrm{Vect}(U) \to \mathrm{Vect}(V)$ which maps vector fields on U to vector fields on V by the formula:

$$f_*(A)(f(x)) = df_x(A(x)), \quad A \in \mathrm{Vect}(U), x \in U. \tag{2.1}$$

Equivalently,

$$f_*(A)(y) = df_{f^{-1}(y)}\left(A(f^{-1}(y))\right), \quad y \in V.$$

Suppose we are given a coordinate system $(x_1, \dots, x_n)$ in U. Then

$$f_*\left(\frac{\partial}{\partial x_j}\right) = \frac{\partial f}{\partial x_j}, \; j = 1, \dots, n,$$

or more precisely

$$f_*\left(\frac{\partial}{\partial x_j}\right)(y) = \frac{\partial f}{\partial x_j}\left(f^{-1}(y)\right), \; y \in V.$$

One can also define the *pull-back* operator $f^* : \mathrm{Vect}(V) \to \mathrm{Vect}(U)$ by the formula $f^* = (f^{-1})^*$. In other words, the pull-back is the push-forward by the inverse diffeomorphism. More explicitly, we can write

$$f^*(B)(x) = d\left(f^{-1}\right)(f(x))(B(f(x)) = (df_x)^{-1}(B(f(x)), \quad B \in \mathrm{Vect}(V), x \in U. \tag{2.2}$$

Let us point out that why for differential form the pull-back operator f^* is defined for an *arbitrary smooth map* f, for the case of vector fields both operators, f_* and f^* are defined only for *diffeomorphisms.*

3 Lie bracket of vector fields

Let $A, B \in \mathrm{Vect}(U)$ be two vector fields on a domain $U \subset \mathbb{R}^n$. As it was shown in 52H, there is a vector field $C \in \mathrm{Vect}(V)$, called the *Lie bracket* of the vector fields A and B and denoted by $C = [A, B]$, which is characterized by the following property: *for any smooth function* $\phi : U \to \mathbb{R}$ *one has*

$$L_C \phi = (L_A L_B - L_B L_A)\phi.$$

A surprising fact here is that though the right-hand side of this equation seems to be the second order differential operator, the left-hand side is the first order operator, so the second derivatives on the right side cancel each other.

Recall that the bracket $[A, B]$ has the following properties

- Lie bracket is a bilinear operation;
- $[A, B] = -[B, A]$ (skew-symmetricity);
- $[[A, B]C] + [[B, C], A] + [[C, A], B] = 0$ (Jacobi identity);
- If $A = \sum\limits_1^n a_j \frac{\partial}{\partial x_j}$ and $B = \sum\limits_1^n b_j \frac{\partial}{\partial x_j}$ then

$$[A, B] = \sum_{i=1}^n \left(\sum_{j=1}^n a_j \frac{\partial b_i}{\partial x_j} - b_j \frac{\partial a_i}{\partial x_j} \right) \frac{\partial}{\partial x_i} . \tag{3.1}$$

In this section we will give a new interpretation of the Lie bracket $[A, B]$.

We define the *Lie derivative* $L_A B$ of the vector field B along the vector field A in a similar way as we defined in Section 1 the Lie derivative of a differential form. Namely,

$$L_A B = \frac{d\,(A^s)^*\,B}{ds}\Big|_{s=0}. \tag{3.2}$$

More explicitly,

$$L_A B(x) = \lim_{s \to 0} \frac{d_{A^s(x)}\,(A^{-s})\,(B(A^s(x)) - B(x)}{s}.$$

Similarly, to Proposition 1.3 we have

Proposition 3.1.

$$L_A B = 0 \iff (A^s)^*\,B \equiv B \ \ \textit{for all} \ \ s \in \mathbb{R}.$$

Proof. We have

$$\left.\frac{d\left(A^{s}\right)^{*}B}{ds}\right|_{s=s_0}=\lim_{s\to 0}\frac{\left(A^{s+s_0}\right)^{*}B-\left(A^{s_0}\right)^{*}B}{s}$$
$$=\lim_{s\to 0}\left(A^{s_0}\right)^{*}\left(\frac{\left(A^{s}\right)^{*}B-B}{s}\right)=\left(A^{s_0}\right)^{*}\left(\lim_{s\to 0}\frac{\left(A^{s}\right)^{*}B-B}{s}\right)$$
$$=\left(A^{s_0}\right)^{*}\left(L_A B\right).$$

Hence, if $L_A B=0$ then $\frac{d(A^s)^*B}{ds}$ for all s and hence $\left(A^{s}\right)^{*}B=\left(A^{0}\right)^{*}B=B$. The converse is obvious. ■

Theorem 3.2. *For any two vector fields $A,B\in\mathrm{Vect}(U)$*

$$L_A B=[A,B].$$

Proof. Note that $A^s(x)=x+sA(x)+o(s)$. Hence, we can write

$$d_y A^{-s}=\mathrm{Id}-s d_y A+o(s),$$

where we view here A as a map $\mathbb{R}^n\to\mathbb{R}^n$. Furthermore, plugging $y=A^s(x)$ we get

$$d_{A^s(x)}A^{-s}=\mathrm{Id}-s d_x A+o(s),$$

because $s(d_y A-d_x A)=o(s)$. We also have $B(A^s(x))=B(x+sA(x)+o(x))=B(x)+s d_x B(A(x))+o(s)$. Thus, ignoring terms $o(s)$-terms we get

$$L_A B=\lim_{s\to 0}\frac{1}{s}\left(d_{A^s(x)}\left(A^{-s}\right)\left(B(A^s(x))\right)-B(x)\right)$$
$$=\lim_{s\to 0}\frac{1}{s}\left(\left(\mathrm{Id}-s d_x A\right)\right)\left(B(x)+s d_x B(A(x))\right)-B(x))$$
$$=\lim_{s\to 0}\frac{1}{s}\left(B(x)-s d_x A(B)+s d_x B(A)-B(x)\right)=d_x B(A)-d_x A(B).$$

But the right-hand-side expression written in coordinates has the form

$$d_x B(A)-d_x A(B)=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}a_j\frac{\partial b_i}{\partial x_j}-b_j\frac{\partial a_i}{\partial x_j}\right)\frac{\partial}{\partial x_i}$$

which coincides with the expression (3.1) for the Lie bracket. ■

Exercise 3.3. Prove that for any smooth function ϕ one has

$$L_{[A,B]}\phi = \frac{\partial^2(\phi \circ A^s \circ B^t)}{\partial s \partial t}.$$

4 First integrals

Suppose we are given a differential equation

$$\dot{x} = A(x), \tag{4.1}$$

where A is a vector field on the domain $U \subset \mathbb{R}^n$ A function $\phi : U \to \mathbb{R}$ is called a *first integral*, or simply an *integral* of equation (4.1) if it is constant on solutions of this equation, or equivalently on integral curves of the vector field A.

Clearly, a necessary and sufficient condition for ϕ to be an integral is to satisfy the equation $L_A\phi = 0$. Here $L_A\phi$ denotes the directional derivative of ϕ along A.

If ϕ is an integral of (3.2) then the solutions are contained in the level sets of the function ϕ, and hence, this allows us to reduce the order of equation by 1. If (3.2) has two integrals ϕ_1, ϕ_2, then the solutions lie inn the intersection of level sets $\{\phi_1 = c_1\}$ and $\{\phi_2 = c_2\}$, $c_1, c_2 \in \mathbb{R}$. Hence, if these level sets transverse to each other (which means that the differential $d\phi_1$ and $d\phi_2$ are linearly independent at every point of the intersection), then the solutions lie in $\{\phi_1 = c_1\} \cap \{\phi_2 = c_2\}$, which allows to further reduce the order of the system. If the order is reduced to 1 then the equation can be explicitly integrated in quadratures. Such systems are called *completely intregrable.*

Some important examples of integrals which come from Mechanics are discussed in the next section.

5 Hamiltonian vector fields

Consider the vector space $\mathbb{R}^{2n}$ with coordinates $(p_1, \dots, p_n, q_1, \dots, q_n)$ and a closed differential 2-form $\omega = \sum\limits_1^n dp_i \wedge dq_i$. We will wrote $p = (p_1, \dots, p_n)$ and $q = q_1, \dots, q_n)$. Given a function

$H : \mathbb{R}^{2n} \to \mathbb{R}$ we will denote

$$\frac{\partial H}{\partial q} := \left(\frac{\partial H}{\partial q_1}, \dots, \frac{\partial H}{\partial q_n}\right), \quad \frac{\partial H}{\partial p} := \left(\frac{\partial H}{\partial p_1}, \dots, \frac{\partial H}{\partial p_n}\right).$$

Note that this form is non-degenerate, i.e. its matrix is non-degenerate at every point. Therefore, the map $J : \mathrm{Vect}(\mathbb{R}^{2n} \to \Omega^1(\mathbb{R}^{2n})$ given by the formula $X \mapsto X \lrcorner \omega$ is an isomorphism between the space $\mathrm{Vect}(\mathbb{R}^{2n}$ of vector fields and the space $\Omega^1(\mathbb{R}^{2n})$ of differential 1-forms on $\mathbb{R}^n$. In coordinates the map J associates with a vector field $(P_1, \dots, P_n, Q_1, \dots, Q_n)$ the differential form $\sum\limits_1^n Q_i dp_i - P_i dq_i$.

Lemma 5.1. *Given a vector field A on $\mathbb{R}^{2n}$ the differential 1-form $J(A) = A \lrcorner \omega$ is closed if and only if $L_A \omega = 0$.*

Proof. Indeed, according to Cartan's formula (1.4) we have $L_A\omega = d(A \lrcorner \omega) = dJ(A)$ because ω is closed. ■

Given a function $H : \mathbb{R}^{2n} \to \mathbb{R}$ we denote by X_H the vector field $-J^{-1}(dH)$. Vector fields obtained by this construction are called *Hamiltonian.*

To find a coordinate expression for X_H we write $X_H = \sum\limits_1^n a_i \frac{\partial}{\partial p_i} + b_i \frac{\partial}{\partial q_i}$. Then

$$X_H \lrcorner \omega = \left(\sum_1^n a_i \frac{\partial}{\partial p_i} + b_i \frac{\partial}{\partial q_i}\right) \lrcorner \sum_1^n dp_i \wedge dq_i = \sum_1^n -b_i dp_i + a_i dq_i.$$

Hence, the equation

$$X_H \lrcorner \omega = -dH = -\sum_1^n \frac{\partial H}{d} p_i + \frac{\partial H}{\partial q_i} dq_i$$

implies $a_i = -\frac{\partial H}{\partial q_i}, b_i = \frac{\partial H}{\partial p_i}$, $i = 1, \dots, n$. Thus,

$$X_H = \sum_1^n -\frac{\partial H}{\partial q_i}\frac{\partial}{\partial p_i} + \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i}.$$

In a shorter form, omitting indices we will write

$$X_H = -\frac{\partial H}{\partial q}\frac{\partial}{\partial p} + \frac{\partial H}{\partial p}\frac{\partial}{\partial q}.$$

Thus the system of differential equations corresponding to the vector field X_H has the form

$$\begin{aligned} \dot{p} &= -\frac{\partial H}{\partial q} \\ \dot{q} &= \frac{\partial H}{\partial p}. \end{aligned} \tag{5.1}$$

These equations play an important role in Mechanics, and called *Hamilton canonical equations.* They describe the phase flow of a mechanical system. Here the coordinates $q = (q_1, \dots, q_n)$ determine a position of the system, or a point in the *configuration space* of the mechanical system. The coordinates $p = (p_1, \dots, p_n)$ are called *momenta* and can be viewed as vectors of the cotangent bundle to the configuration space. The function H is the full energy of the system expressed through coordinates and momenta.

Lemma 5.2. *The function H is a first integral of the equation* (5.1), *i.e.* $L_{X_H} H = 0$.

Proof.

$$L_{X_H} H = dH(X_H) = -\frac{\partial H}{\partial p}\frac{\partial H}{\partial q} + \frac{\partial H}{\partial q}\frac{\partial H}{\partial p} = 0.$$

■

Example 5.3. Consider Newton equations

$$\ddot{q}_i = -\frac{\partial U}{\partial q_i}, i = 1, \dots, n,$$

or in shorter notation

$$\ddot{q} = -\frac{\partial U}{\partial q} = -\nabla U.$$

Reducing it to a system of first order equation we get

$$\dot{p} = -\frac{\partial U}{\partial q} \tag{5.2}$$

$$\dot{q} = p. \tag{5.3}$$

Consider the *full energy* $H(p, q) = \sum\limits_1^n \frac{p_i^2}{2} + U(q) = \frac{1}{2}p^2 + U(q)$. Then $\frac{\partial H}{\partial q} = \frac{\partial U}{\partial q}$ and $\frac{\partial H}{\partial p} = p$, and hence equation (5.2) takes the form (5.1) with this Hamiltonian function H. Lemma 5.2 is the law of conservation law of energy.

Lemma 5.4. *Let X_H be a Hamiltonian vector field and X_H^s, the phase flow it generates. Then $(X_H^s)^* \omega = \omega$ for all $s \in \mathbb{R}$. In other words, the flow of a Hamiltonian vector field preserves the form ω.*

Proof. It is sufficient to prove that $L_{X_H}\omega = 0$. Using Theorem 1.2 we get

$$L_{X_H}\omega = d(X_H \lrcorner\, \omega) + X_H \lrcorner\, d\omega.$$

But ω is closed, and hence $d\omega = 0$, while $X_H \lrcorner\, \omega = dH$. Thus, $L_{X_H}\omega = ddH = 0$. ■

6 Canonical transformations

The equations (5.1) are called canonical because they are invariant with respect to a large group of transformation of the phase space. Let us call a diffeomorphism $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ a *symplectomorphism* (or alternatively a *canonical transformation*) if it preserves the form ω. Then it preserves also the form of the equations (5.1). Indeed, suppose $f(p, q) = (\widetilde{p}, \widetilde{q})$. Then $f^*(\omega) = f^*(dp \wedge dq) = d\widetilde{p} \wedge d\widetilde{q} = \omega = dp \wedge dq$. Thus if we express the function $H(p, q)$ through the coordinates $\widetilde{p}, \widetilde{q}$, $H(p, q) = \widetilde{H}(\widetilde{p}, \widetilde{q})$ then the equation (5.1) will take the same form in coordinates $(\widetilde{p}, \widetilde{q})$:

$$\begin{aligned} \dot{\widetilde{p}} &= -\frac{\partial \widetilde{H}}{\partial \widetilde{q}} \\ \dot{\widetilde{q}} &= \frac{\partial \widetilde{H}}{\partial \widetilde{p}}. \end{aligned} \tag{6.1}$$

The following proposition provides an important class of canonical transformations,

Proposition 6.1. *Consider any diffeomorphism $f : U \to V$ between two domains $U, V \subset \mathbb{R}^{2n}$. Let Df be the Jacobi matrix of the map U. Then the map*

$$(p, q) \mapsto \left((Df)^{-1} \right)^T p, f(q))$$

is a symplectomorphism $\widehat{f}$ of the domain $\widehat{U} = \{p \in \mathbb{R}^n, q \in U\}$ to the domain $\widehat{V} = \{p \in \mathbb{R}^n, q \in V\}$. Here $\left((Df)^{-1} \right)^T$ is the matrix transpose to inverse of the Jacobi matrix Df.

In other words, *any change of q-coordinates extends to a canonical change of the (p, q)-coordinates.*

Proof. Let us denote the elements of the matrix $(Df)^{-1}$ by g_{ij}, $i, j = 1, \ldots, n$. Thus, $\sum_i^n g_{ji} \frac{\partial f_i}{\partial q_k} = \delta_{jk}$, $\delta_{jk} = 1$ if $j = k$ and $\delta_{jk} = 0$ if $j \neq k$.

Let us compute $\widehat{f}^*(pdq) = \widehat{f}^* \left(\sum_1^n p_i dq_i \right)$. We have

$$\widehat{f}(p_1, \ldots, p_n, q_1, \ldots, q_n) = \left(\sum_1^n g_{j1} p_j, \ldots, \sum_1^n g_{jn} p_j, f_1(q), \ldots, f_n(q) \right).$$

Hence,

$$\begin{aligned}
\widehat{f}^*(pdq) = \widehat{f}^* \left(\sum_1^n p_i dq_i \right) &= \sum_{i=1}^n \sum_{j=1}^n g_{ji} p_j df_i \\
&= \sum_{i,j,k=1}^n g_{ji} \frac{\partial f_i}{\partial q_k} p_j dq_k = \sum_{j,k=1}^n \delta_{jk} p_j dq_k \\
&= \sum_1^n p_k dq_k = pdq.
\end{aligned}$$

Hence,

$$\widehat{f}^* \omega = \widehat{f}^* dp \wedge dq = d(\widehat{f}^*(pdq)) = d(pdq) = dp \wedge dq = \omega.$$

■

Corollary 6.2. . *Suppose that there exists a change of coordinates $\widetilde{q} = f(q)$ such that in new coordinates the Hamiltonian function H is independent of the coordinate $\widetilde{q}_1$. Then $\widetilde{p}_1 = \sum_1^n g_{j1} p_j$ is a first integral of the system* (5.1)*. Here the notation g_{ij} stands for the elements of the matrix $(Df)^{-1}$.*

Proof. Let us extend the coordinate change $q \mapsto \widetilde{q} = f(q)$ to a canonical change of coordinates $(p, q) \mapsto (\widetilde{p}, \widetilde{q}) = \widetilde{f}(p, q)$ as in Proposition 6.1. Then the equation in the new coordinates $(\widetilde{p}, \widetilde{q})$ also has the canonical Hamiltonian form (6.1). Then $\dot{\widetilde{p}}_1 = \frac{\partial H}{\partial \widetilde{q}_1} = 0$ because by assumption the Hamiltonian is independent of the coordinate $\widetilde{q}_1$. Hence $\widetilde{p}_1 = \sum_1^n g_{j1} p_j$ is constant along trajectories, i.e. it is a first integral. ■

7 Example: angular momentum

Consider a Newton equation

$$\ddot{q} = -\nabla U(q), \quad q \in \mathbb{R}^3, \tag{7.1}$$

which describes the motion of a particle of mass 1 in a field with a potential energy function $U(q)$. Suppose there exists an axis l in $\mathbb{R}^3$ such that the function $U(q)$ remains invariant with respect to rotations around l.

The system (7.1) can be rewritten in the Hamiltonian form (5.1) with the Hamiltonian function $H = \frac{p^2}{2} + U(q) = \frac{p_1^2}{2} + \frac{p_2^2}{2} + \frac{p_3^2}{2} + U(q_1, q_2, q_3)$. Let us assume for simplicity that the q_3-axis coincides with the axis l.

Let us change coordinates (q_1, q_2, q_3) to cylindrical coordinates (ϕ, r, z):

$$q_1 = r\cos\phi, q_2 = r\sin\phi, q_3 = z.$$

Equivalently,

$$\phi = \arctan\frac{q_2}{q_1}, r = \sqrt{q_1^2 + q_2^2}, z = q_3.$$

Computing the Jacobi matrix $\frac{D(\phi,r,z)}{D(q_1,q_2,q_3)}$ we get

$$\begin{pmatrix} \frac{\partial\phi}{\partial q_1} & \frac{\partial\phi}{\partial q_2} & \frac{\partial\phi}{\partial q_3} \\ \frac{\partial r}{\partial q_1} & \frac{\partial r}{\partial q_2} & \frac{\partial r}{\partial q_3} \\ \frac{\partial z}{\partial q_1} & \frac{\partial z}{\partial q_2} & \frac{\partial z}{\partial q_3} \end{pmatrix} = \begin{pmatrix} -\frac{q_2}{q_1^2+q_2^2} & \frac{q_1}{q_1^2+q_2^2} & 0 \\ \frac{q_1}{\sqrt{q_1^2+q_2^2}} & \frac{q_2}{\sqrt{q_1^2+q_2^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then the inverse matrix is equal to

$$\begin{pmatrix} -q_2 & \frac{q_1}{\sqrt{q_1^2+q_2^2}} & 0 \\ q_1 & \frac{q_2}{\sqrt{q_1^2+q_2^2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

[2] Let us extend the coordinate change $(r, \phi, z) \mapsto (q_1, q_2, q_3)$ to a canonical coordinate change

$$(q_1, q_2, q_3, p_1, p_2, p_3) \mapsto (\phi, r, q_3, p_\phi, p_r, p_z),$$

where we denoted by p_r, p_ϕ, p_z momenta variables corresponding to new coordinates (r, ϕ, z). In fact, we need only the coordinate p_ϕ which is given by $p_\phi = -p_1 q_2 + q_1 p_2$. Thus, the function $-p_1 q_2 + p_2 q_1$ is the first integral. It is called *the angular momentum around the* q_3*-axis.*

---

[2]Of course, in this case it would be easier to compute the Jacobi matrix of the map $(\phi, r, z) \to (q_1, q_2, q_3)$ and then change the coordinates in the result. However, we are following here precisely the scheme provided by Proposition 6.1