# HAMILTON'S EQUATIONS AND CANONICAL TRANSFORMATIONS
# October 3, 2024

ARTHUR JAFFE

Hamilton's formulation of classical mechanics brought to the fore the motion of a point in *phase space*. One point $\xi$ in the phase space for a particle is determined by both position $\vec{q}$ and a momentum $\vec{p}$. For a particle in 3-dimensions, this means that the phase space is 6 dimensional. So for $n$ particles in 3-space, the phase space is $6n$ dimensional. Particle motion in phase space is determined by Hamilton's equations for $\xi(t)$. They are first order differential equations in time, and a trajectory $\xi(t)$ in phase space is a solution to these equations given an initial condition $\xi(t_0)$ at the initial time; one does not require a condition on the initial velocity.

Hamiltonian mechanics also gives new insights into the relation between symmetry and conservations laws. It presents a framework for which one can imagine how to go in both directions

$$\text{Symmetry} \Longleftrightarrow \text{Conservation Law}\ ,$$

while Noether's theorem only demonstrates going from a symmetry to a conserved quantity, the Noether charge.

Furthermore, Hamiltonian mechanics gives a natural way to consider Noether charges as generators of *infinitesimal* symmetries. This relates a Noether change to an element of a Lie algebra, which generates a global symmetry described by a Lie group. These beautiful ideas will emerge here in the study of classical mechanics. But their generalizations pervade much of modern physics and mathematics. The algebraic and group theoretical aspects of Hamiltonian mechanics should provide insights into all your other more advanced courses.

## 1. Hamilton's Equations on Phase Space

Although Hamilton's equations are a consequence of Lagrange's equations, there were huge gaps in time during the evolution of classical mechanics. The period from the birth of Euler to the birth of Hamilton spans almost 100 years. Euler (1707–1783) studied the calculus of variations. He was 30 years older than Lagrange (1736–1813), who introduced the Lagrangian approach in 1788. Some 50 years later, Hamilton (1805–1865), worked on mechanics during an early period of his life and found his equations in 1833. This was already 20 years after Lagrange had died.

Hamilton was an undergraduate at Trinity College, Dublin. While only 23 years old, he was appointed Professor. Eventually he became Astronomer Royale of Ireland and lived in the observatory estate, located in Dunsink. He sometimes walked along the river to Dublin. On one of these walks, he worked out in his mind the quarternion algebra, and inscribed it with a stone on Broome Bridge.

The Hamilton equations of motion for a system of $N$ configuration-space coordinates $q$ and the $N$ momentum-space coordinates $p$ take place in the $2N$-dimensional space of coordinates $(q, p)$, known as phase space. They are the $2N$ first-order differential equations

$$\boxed{\dot{q}_j = \frac{\partial H}{\partial p_j}\ , \quad \text{and} \quad \dot{p}_j = -\frac{\partial H}{\partial q_j}\ , \ \text{for}\ j = 1, \ldots, N}\ . \tag{1.1}$$

FIGURE 1. Broome Bridge, March 2008.

As the equations are first order, they require only that we give the initial values at one time $t_0$, namely $q(t_0), p(t_0)$.

If one starts from a Lagrangian system, one defines the Hamiltonian as

$$H(q,p,t) = \sum_{k=1}^{N} p_k \dot{q}_k - \mathcal{L}(q,\dot{q},t) \ , \tag{1.2}$$

using the relation between $p, q$, and $\dot{q}$

$$p_k = \frac{\partial \mathcal{L}(q,\dot{q},t)}{\partial \dot{q}_k} \ , \tag{1.3}$$


FIGURE 2. Hamilton Placard on Broome Bridge.

to eliminate $\dot{q}$ from $H$, and to express it in terms of $q$ and $p$. In other words, we solve the equation (1.3) for $\dot{q}_j =$ A major nice thing about Hamilton equations, compared with Lagrange equations, is that one can easily imagine a route to connect symmetries with conservation laws in both directions:

$$\boxed{\text{Conservation Law} \Longleftrightarrow \text{Continuous Symmetry}}$$

Noether's theorem gave a route to go from a continuous symmetry to a conserved quantity, but not backwards. In this section we develop the ideas of how to go in both directions. The conceptual basis of this relation is beautiful and simple, and we attempt to make them clear. However there are technical obstacles to prevent the simple implementation of these ideas and to find a complete correspondence back and forth between conservation laws and continuous symmetry. Some of these obstacles are not yet completely understood. They focus on showing in certain instances (which we will not study) that the Hamilton's equations do have solutions.

1.1. **Equation on Phase Space.** Let us first comment on a subtlety about the partial derivatives. In Lagrangian mechanics we consider the variables $q$ and $\dot{q}$ to be the independent variables. Whenever the momentum $p$ occurs, we regard $p = p(q, \dot{q})$ as a function of $q$ and $\dot{q}$. Here $q$ is called *configuration space* while $\dot{q}$ is called *velocity space*. In other words, in Lagrange's equations (1.5), the expression $\frac{\partial \mathcal{L}}{\partial \dot{q}_j}$ means

$$\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_j}\right)_{q_i, \dot{q}_\ell, \text{ for all } i, \text{ and for } \ell \neq j} . \tag{1.4}$$

Here the subscripts indicate the variables held fixed in differentiation. However, in order to simplify notation, we do not write each time which variables are independent–unless there is a real possibility of confusion. When we consider the Lagrange equations for $\mathcal{L}(q, \dot{q})$ we always assume that $q$, with components $q_i$, and $\dot{q}$, with components $\dot{q}_j$, are the independent variables.

On the other hand, in Hamiltonian mechanics the variables $q$ and $p$ are the basic independent variables, and $(q, p)$ is said to be a point in phase space. Whenever the velocity $\dot{q}$ occurs, it is regarded as a function $\dot{q} = \dot{q}(q, p)$ of $q$ and $p$. We will derive Hamilton's equations from Lagrange's equations by assuming that $q$ and $p$ are the independent variables, and working out the consequences. Thus if we express $H$ as in (1.2), then the derivative of the second term is

$$\left(\frac{\partial \mathcal{L}}{\partial p_j}\right)_{q_i, p_\ell, \text{ for all } i, \text{ and for } \ell \neq j} = \sum_{k=1}^{N} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_k}\right)_{q_i, \dot{q}_\ell, \text{ for all } i, \text{ and for } \ell \neq k} \left(\frac{\partial \dot{q}_k}{\partial p_j}\right)_{q_i, p_{\ell'}, \text{ for all } i, \text{ and for } \ell' \neq j} .$$

While this can be confusing, one must decide which variables are independent and which variables are dependent; and one must stick to it! The choice will depend on the context. In the derivation of Hamilton's equations below, we take $q$ and $p$ to be independent, although we assume Lagrange's equations, which are derived assuming that $q$ and $\dot{q}$ are independent.

On the other hand, we could reverse the argument and derive Lagrange's equations from Hamilton's equations.

## 1.2. **Equivalence of Hamilton's Equations and Lagrange's Equations.**

**Proposition 1.1.** *The Lagrange's equations hold for a trajectory $q(t)$, namely*

$$p_j = \frac{\partial \mathcal{L}(q,\dot{q},t)}{\partial \dot{q}_j}\ , \quad \textit{satisfies} \quad \dot{p}_j = \frac{d}{dt}\frac{\partial \mathcal{L}(q,\dot{q},t)}{\partial \dot{q}_j} = \frac{\partial \mathcal{L}(q,\dot{q},t)}{\partial q_j(t)}\ , \quad \textit{for} \quad j=1,\dots,N\ , \tag{1.5}$$

*if and only if the Hamilton equations hold on phase space, namely*

$$\dot{q}_j = \frac{\partial H(q,p,t)}{\partial p_j}\ , \quad \textit{and} \quad \dot{p}_j = -\frac{\partial H(q,p,t)}{\partial q_j}\ , \ \textit{for}\ j=1,\dots,N\ . \tag{1.6}$$

*Proof.* The entire question here revolves around what is the meaning of the partial derivatives. In Lagrangian mechanics, one expresses the Lagrangian $\mathcal{L}(q,\dot{q},t)$ in the variables given by the coordinates $q$ and velocities $\dot{q}$. So a partial derivative $\frac{\partial}{\partial q_i}$ means vary $q_i$, while keeping fixed both all the $\dot{q}_j$'s as well as those $q_j$'s with $j \neq i$. Similarly $\frac{\partial}{\partial \dot{q}_i}$ means vary $\dot{q}_i$, while keeping fixed both all the $q_j$'s as well as those $\dot{q}_j$'s with $j \neq i$. On the other hand, in Hamiltonian mechanics one expresses the Hamiltonian $H(q,p,t)$ as a function of the coordinates $q$ and the momenta $p$, defined as $p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$. So in Hamiltonian mechanics the partial derivative $\frac{\partial}{\partial q_i}$ means vary $q_i$, while keeping fixed both all the $p_j$'s as well as those $q_j$'s with $j \neq i$. Similarly $\frac{\partial}{\partial p_i}$ means vary $p_i$, while keeping fixed both all the $q_j$'s as well as those $p_j$'s with $j \neq i$.

With this in mind, and with $H(q,p,t) = \sum_{j=1}^N p_j\dot{q}_j - \mathcal{L}(q,\dot{q},t)$, let us calculate the derivative of $H(q,p,t)$ with respect to $p_j$, under the assumptions explained above. Thus we keep $p_i$ fixed for $i \neq j$, and keep $q_i$ fixed for all $i$. This yields

$$\begin{aligned}
\frac{\partial H}{\partial p_j} &= \frac{\partial}{\partial p_j}\left(\sum_{j=1}^N p_j\dot{q}_j - \mathcal{L}(q,\dot{q},t)\right) \\
&= \dot{q}_j + \sum_{k=1}^N \left(p_k\frac{\partial \dot{q}_k}{\partial p_j} - \frac{\partial \mathcal{L}}{\partial \dot{q}_k}\frac{\partial \dot{q}_k}{\partial p_j}\right) \\
&= \dot{q}_j + \sum_{k=1}^N \left(p_k\frac{\partial \dot{q}_k}{\partial p_j} - p_k\frac{\partial \dot{q}_k}{\partial p_j}\right) = \dot{q}_j\ .
\end{aligned}$$

The first term $\dot{q}_j$, as well as the first sum, arise from differentiating $\sum_k p_k\dot{q}_k$. The other terms come from differentiating $\mathcal{L}$. As $p_k = \frac{\partial \mathcal{L}}{\partial \dot{q}_k}$ the terms that are in parentheses in (1.7) cancel, and the first $N$ Hamilton equations hold, namely those for $\dot{q}_j$ in (1.6).

Similarly, one can differentiate (1.7) with respect to $q_j$, keeping $q_i$ fixed for $i \neq j$, and keeping $p_i$ fixed for all $i$. Assuming the Lagrange equations, one obtains

$$\begin{aligned}
\frac{\partial H}{\partial q_j} &= \sum_{k=1}^{N} p_k \frac{\partial \dot{q}_k}{\partial q_j} - \left( \frac{\partial \mathcal{L}}{\partial q_j} + \sum_{k=1}^{N} \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \frac{\partial \dot{q}_k}{\partial q_j} \right) \\
&= \sum_{k=1}^{N} \left( p_k \frac{\partial \dot{q}_k}{\partial q_j} - \frac{\partial \mathcal{L}}{\partial \dot{q}_k} \frac{\partial \dot{q}_k}{\partial q_j} \right) - \frac{\partial \mathcal{L}}{\partial q_j} \\
&= -\frac{\partial \mathcal{L}}{\partial q_j} = -\dot{p}_j \; . \qquad (1.7)
\end{aligned}$$

In the middle line we use $p_k = \frac{\partial \mathcal{L}}{\partial \dot{q}_k}$ to exhibit cancellation of the terms in parentheses. In the last line, we use Lagrange equations to set

$$\frac{\partial \mathcal{L}}{\partial q_j} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_j} = \dot{p}_j \; . \qquad (1.8)$$

This completes the proof of Hamilton's equations from Lagrange's equations.

In order to derive Lagrange's equations from Hamilton's equations, the procedure is similar. We first derive the formula for the momentum, based on reversing the identity relating $H$ to $\mathcal{L}$, namely writing $\mathcal{L} = (\sum_i p_i \dot{q}_i) - H$. Then if we differentiate with respect to $\dot{q}_j$, keeping all the $q_i$ fixed, as well as the $\dot{q}_j$ for $j \neq i$, then

$$\begin{aligned}
\frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial \dot{q}_j} &= p_j + \sum_i \left( \dot{q}_i \frac{\partial p_i}{\partial \dot{q}_i} \right) - \frac{\partial H}{\partial \dot{q}_j} \\
&= p_j + \sum_i \left( \dot{q}_i \frac{\partial p_i}{\partial \dot{q}_i} - \frac{\partial H}{\partial q_i} \frac{\partial q_i}{\partial \dot{q}_i} - \frac{\partial H}{\partial p_i} \frac{\partial p_i}{\partial \dot{q}_i} \right)
\end{aligned}$$

In the parentheses, the second term vanishes since $q$ and $\dot{q}$ are independent variables, so $\frac{\partial q_i}{\partial \dot{q}_i} = 0$. In the last term, we use the Hamilton equation $\frac{\partial H}{\partial p_i} = \dot{q}_i$. Thus the terms in the parenthesis vanish and

$$\frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial \dot{q}_j} = p_j \; .$$

We can now continue to derive the Lagrange equations. Let us calculate $\frac{\partial \mathcal{L}}{\partial q_j}$. keeping the other $q$'s for $i \neq j$ as well as all the $\dot{q}$'s fixed. We use Hamilton's equations to obtain

$$\begin{aligned}
\frac{\partial \mathcal{L}(q, \dot{q}, t)}{\partial q_j} &= \frac{\partial}{\partial q_j} \left( \sum_i p_i \dot{q}_i - H \right) \\
&= \sum_i \left( \frac{\partial p_i}{\partial q_j} \dot{q}_i - \frac{\partial H}{\partial p_i} \frac{\partial p_i}{\partial q_j} \right) - \frac{\partial H}{\partial q_j} \\
&= \dot{p}_j \; .
\end{aligned}$$

The Hamilton equation $\dot{q}_i = \frac{\partial H}{\partial p_i}$ shows that the terms in the middle line that are in parenthesis cancel. On the other hand the Hamilton equation $\frac{\partial H}{\partial q_j} = \dot{p}_j$ yields the Lagrange equation as claimed. □

1.3. **Hamilton's Equations as a Phase Space Flow.** It sometimes simplifies things to put $q$ and $p$ together into one $2N$-dimensional vector. This is called a point in *phase space*, and denote this point by a $2N$ dimensional vector

$$\xi = \begin{pmatrix} q \\ p \end{pmatrix} , \tag{1.9}$$

where $q$ and $p$ are $N$-dimensional points in the space of coordinates and momentum.

The simplest phase space occurs for a system like a simple harmonic oscillator, defined by coordinates $q$ and momenta $p$ in $\mathbb{R}^N$, so the phase space is $\xi \in \mathbb{R}^{2N}$. In this case the phase space is actually a linear vector space: the linear combination of two points in phase space are another point in phase space. The actual motion of a point in phase space may lie on a subspace that is not a vector space. But it is often useful to regard the manifold on which particles move as embedded in a larger vector space.

We label the components of $\xi$ as

$$\xi = \begin{pmatrix} q \\ p \end{pmatrix} , \quad \text{with components } \xi_i = \begin{cases} q_i & \text{for } i = 1, \ldots, N \\ p_{i-N} & \text{for } i = N+1, \ldots, 2N \end{cases} . \tag{1.10}$$

Written on phase space Hamilton's equations have the form

$$\frac{d}{dt} \begin{pmatrix} q \\ p \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \nabla_q H(q,p,t) \\ \nabla_p H(q,p,t) \end{pmatrix} .$$

Since the matrix coefficient comes up often, we give a name to the $2N \times 2N$ matrix

$$\Gamma = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} , \tag{1.11}$$

made up as a $2 \times 2$ array of $N \times N$ matrices $0$ or $\pm I$. Also

$$\Gamma^2 = -I , \quad \text{and} \quad \det \Gamma = 1 . \tag{1.12}$$

To see the latter, note that $\Sigma = -i\Gamma$ is hermitian and has square $I$. So $\Sigma$ has an orthonormal basis of eigenvectors, with eigenvalues $\pm 1$. The sum of these eigenvalues equals the trace of $\Sigma$, which is zero, so there must be an equal number $N$ of eigenvalues $1$ and eigenvalues $-1$. The product of these eigenvalues equals $\det \Sigma = (-1)^N$. Hence $\det \Gamma = (-1)^N \det(iI) = (-1)^N i^{2N} = 1$. The phase-space notation lets one summarize Hamilton's equations in a compact form,

$$\boxed{\dot{\xi} = \Gamma\, \nabla_\xi H(\xi)} . \tag{1.13}$$

## 2. Coordinate Change and Poisson Brackets

2.1. **Poisson Brackets.** A point in phase space puts together the coordinate $q$ and the canonically conjugate momentum $p$ into a single variable $\xi = (q, p)$, namely a point in phase space. Thus functions like angular momentum, the energy, or the Lenz vector are functions of the phase space coordinate. With $N$ degrees of freedom, the phase space is $2N$ dimensional.

**Definition 2.1 (Poisson Bracket).** *Let $A, B$ denote two differentiable functions on phase space of dimension $2N$. Denote their Poisson bracket by $[A\,,\,B]_\xi$, where*

$$\boxed{[A\,,\,B]_\xi = \sum_{i=1}^{N} \left( \frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right)}\,. \tag{2.1}$$

We often use a shorthand and write

$$\boxed{[A\,,\,B]_\xi = \sum_{i,j=1}^{2N} \frac{\partial A}{\partial \xi_i}\, \Gamma_{ij}\, \frac{\partial B}{\partial \xi_j}\,,} \tag{2.2}$$

where $\Gamma$ is the $2N \times 2N$ matrix (1.11) used above.

2.2. **The Poisson Bracket as a Derivative (Product Rule).** The Poisson bracket $[A\,,\,B]_\xi$ is linear in both $A$ and $B$. So for fixed $A$, the Poisson bracket $[A\,,\,B]_\xi$ is a linear transformation of functions $B$ on phase space, which takes the function $B(\xi)$ into a new function $[A\,,\,B]_\xi$. In fact this new function is a derivative of $B$, and we can denote that transformation by $D_A$, as it depends on $A$. Let

$$D_A = \sum_{i,j=1}^{2N} \frac{\partial A(\xi)}{\partial \xi_i}\, \Gamma_{ij}\, \frac{\partial}{\partial \xi_j}\ . \tag{2.3}$$

With this convention,

$$D_A\, B = [A\,,\,B]_\xi\ . \tag{2.4}$$

This entails the usual *product rule* for derivatives of a product. One can express this in two forms: one as a derivative, or another as a Poisson bracket,

$$\boxed{D_A(BC) = (D_A B)C + B D_A C}\,, \quad \text{or} \quad \boxed{[A\,,\,BC]_\xi = [A\,,\,B]_\xi\, C + B\,[A\,,\,C]_\xi}\,. \tag{2.5}$$

2.3. **Hamilton Equations as Poisson Brackets.** One could also write Hamilton's equations for a trajectory $\xi(t)$ in phase space as

$$\dot{\xi}_j(t) = -\,[H\,,\,\xi_j(t)]_\xi\ . \tag{2.6}$$

Alternatively, Hamilton's equations are

$$\dot{\xi}_j(t) = -D_H\, \xi_j(t)\ . \tag{2.7}$$

2.4. **Fundamental Poisson Brackets.** The Poisson brackets of two coordinates on phase space play a fundamental role. In terms of the $q$ and $p$ variables they are,

$$\boxed{[q_i\,,\,q_j]_\xi = 0\;,\quad [p_i\,,\,p_j]_\xi = 0\;,\quad \text{and}\quad [q_i\,,\,p_j]_\xi = \delta_{ij}\;,\quad \text{for } i,j = 1,\ldots,N}\,. \tag{2.8}$$

These are called the *fundamental Poisson brackets.* In terms of the coordinates $\xi$, one can write the equivalent expression,

$$\boxed{[\xi_i\,,\,\xi_j]_\xi = \Gamma_{ij}\;,\;\; \text{for}\;\; i,j = 1,\ldots,2N}\,. \tag{2.9}$$

2.5. **Composition of Jacobians.** Recall the definition of the Jacobian matrix for the change of coordinates $\xi \mapsto \Xi$ on phase space:

$$J_{\Xi\leftarrow\xi} = J_{\xi\to\Xi} = \frac{\partial \Xi}{\partial \xi} \quad \text{has components} \quad J_{ij} = \frac{\partial \Xi_j}{\partial \xi_i}\;. \tag{2.10}$$

The inverse transformation is $\Xi \mapsto \xi$ and has the Jacobian

$$J_{\xi\leftarrow\Xi} = J^{-1}_{\Xi\leftarrow\xi} \;\; = J_{\Xi\to\xi};. \tag{2.11}$$

More generally, recall the composition of Jacobians associated with two coordinate changes on phase space is given by matrix multiplication of the Jacobians.

**Proposition 2.2.** *For the transformations*

$$\xi \longrightarrow \Upsilon \longrightarrow \Xi \;\;, \tag{2.12}$$

*one has*

$$J_{\xi\to\Xi} = J_{\xi\to\Upsilon} J_{\Upsilon\to\Xi}\;. \tag{2.13}$$

*Proof.* This is a consequance of

$$(J_{\xi\to\Xi})_{ij} = \frac{\partial \Xi_j}{\partial \xi_i} = \sum_\ell \frac{\partial \Xi_j}{\partial \Upsilon_\ell}\frac{\Upsilon_\ell}{\partial \xi_i} = \sum_\ell (J_{\xi\to\Upsilon})_{i\ell}(J_{\Upsilon\to\Xi})_{\ell j}\;. \tag{2.14}$$

□

This can also be written as a "commutative diagram,"

$$\begin{tikzcd} \xi \arrow[r, "J_{\xi\to\Upsilon}"] \arrow[rr, bend right, "J_{\xi\to\Xi}"'] & \Upsilon \arrow[r, "J_{\Upsilon\to\Xi}"] & \Xi\,. \end{tikzcd} \tag{2.15}$$

2.6. **Exponential Transformations and Jacobians.** In this section we consider a transformation $T = T$ on phase space that is not-necessarily linear. We let $\xi_j \to T\xi_j$ define the transformation $\xi \to T\xi$. The Jacobian matrix $J$ of $T$ has matrix elements $J_{ij} = \frac{\partial \Xi_j}{\partial \xi_i}$. The entries of the Jacobian matrix may depend on $\xi$, so that they are no longer necessarily constant.

**Proposition 2.3.** *Let $T$ be a transformation on phase space with Jacobian $J_{\xi\to T\xi}$ and let $\lambda \in \mathbb{C}$. Assume that the transformation $e^{-\lambda T}$ can be defined as a power series in $\lambda$.*[1] *Then it's Jacobian is*

$$J_{\xi\to e^{-\lambda T}\xi} = e^{-\lambda J_{\xi\to T\xi}} \ . \tag{2.16}$$

*Proof.* The Jacobian matrix of $-\lambda T$ is $-\lambda J_{\xi\to T\xi}$. Even if $T, S$ are nonlinear, the Jacobian matrix of $T+S$ is $J_T + J_S$, where we write $J_T$ for $J_{\xi\to T\xi}$. Furthermore, we saw in (2.13) that the Jacobian matrix of $TS$ is $J_T J_S$. Hence, if we can expand the transformation $e^{-\lambda T}$ in a power series, the formula (2.16) follows. □

2.7. **Jacobians and the Solution to Hamilton-Type Equations.** In this section, we consider the Hamilton-type equations on phase space, in which we replace the time parameter $t$ by a general parameter $\lambda$, which may be time or may have some other interpretation. We also replace the Hamiltonian $H = H(\xi)$ by some other function $G(\xi)$ that is an arbitrary function on phase space. It may be an energy, or some other function. These equations we study describe the evolution of the coordinates $\xi_i = \xi_i(\lambda)$ as a function of $\lambda$. They are

$$\boxed{\frac{d}{d\lambda}\,\xi_i(\lambda) = -\sum_{\alpha,\beta} \frac{\partial G(\xi(\lambda))}{\partial \xi_\alpha(\lambda)} \Gamma_{\alpha\beta} \frac{\partial \xi_i(\lambda)}{\partial \xi_\beta(\lambda)} = -D_{G(\xi(\lambda))}\,\xi_i(\lambda)} \ . \tag{2.17}$$

In case $\lambda = t$ and $G = H$, these equations are just Hamilton's equations for the evolution in time of a point in phase space, determined by the Hamiltonian $H = H(\xi)$.

We now study

$$D_{G(\xi(\lambda))}\,C(\xi_i(\lambda)) = \left[G(\xi(\lambda))\,,\,C(\xi(\lambda))\right]_{\xi(\lambda)} = \sum_{\alpha,\beta} \frac{\partial G(\xi(\lambda))}{\partial \xi_\alpha(\lambda)} \Gamma_{\alpha\beta} \frac{\partial C(\xi_i(\lambda))}{\partial \xi_\beta(\lambda)} \ .$$

**Theorem 2.4.** *Let $G(\xi)$ not depend explicitly on $\lambda$, and let $\xi(\lambda)$ satisfies the equation* (2.17)*. Then a function $C(\xi(\lambda))$ on phase space satisfies*

$$\frac{d}{d\lambda}\,C(\xi_i(\lambda)) = -D_{G(\xi(\lambda))}\,C(\xi_i(\lambda)) \ .$$

*In particular, $\frac{dG(\xi(\lambda))}{d\lambda} = 0$, namely $G(\xi)$ is conserved under the evolution of* (2.17)*.*

*Proof.*

$$\begin{aligned}
\frac{dC(\xi(\lambda))}{d\lambda} &= \sum_i \frac{\partial C(\xi(\lambda))}{\partial \xi_i(\lambda)} \frac{d\xi_i(\lambda)}{d\lambda} = -\sum_{i,\alpha,\beta} \frac{\partial C(\xi(\lambda))}{\partial \xi_i(\lambda)} \frac{\partial G(\xi(\lambda))}{\partial \xi_\alpha(\lambda)} \Gamma_{\alpha\beta} \frac{\partial \xi_i(\lambda)}{\partial \xi_\beta(\lambda)} \\
&= -\sum_{i,\alpha,\beta} \frac{\partial C(\xi(\lambda))}{\partial \xi_i(\lambda)} \frac{\partial G(\xi(\lambda))}{\partial \xi_\alpha(\lambda)} \Gamma_{\alpha i} = -\left[G(\xi(\lambda))\,,\,C(\xi(\lambda))\right]_{\xi(\lambda)} \\
&= -D_{G(\xi(\lambda))}\,C(\xi_i(\lambda)) \ .
\end{aligned}$$

Furthermore, taking $C = G$ shows that $G$ is conserved. □

[1] We do not discuss here results in the more complicated case when the power series in $\lambda$ does not converge, but nevertheless $e^{-\lambda T}$ can be defined by some method other than by a power series in $\lambda$.

Transformations on phase space have a natural multiplication, given by the *composition* of the two transformations. If $\Xi$ and $\Upsilon$ are the transformations

$$\xi \mapsto \Xi(\xi)\ , \quad \text{and} \quad \Xi \mapsto \Upsilon(\Xi)\ ,$$

then the composition product $\Upsilon \circ \Xi$ is defined as the transformation

$$\xi \mapsto \Upsilon \circ \Xi\ , \quad \text{namely} \quad (\Upsilon \circ \Xi)(\xi) = \Upsilon(\Xi(\xi))\ . \tag{2.18}$$

**Proposition 2.5.** *The Jacobian matrix of the transformation $\xi \mapsto \Upsilon \circ \Xi$ is the product of the Jacobian matrices for the transformations in the composition in reverse order, namely*

$$\frac{\partial(\Upsilon \circ \Xi)}{\partial \xi} = \frac{\partial \Xi}{\partial \xi}\frac{\partial \Upsilon}{\partial \Xi}\ . \tag{2.19}$$

*Furthermore, the Jacobian matrix of the sum of two transformations is the sum of the Jacobian matrices of each thansformation.*

*Proof.* Let us calculate the Jacobian matrix of the composition. One finds,

$$\begin{aligned}\left(\frac{\partial \Upsilon \circ \Xi}{\partial \xi}\right)_{ij} &= \frac{\partial(\Upsilon \circ \Xi)_j(\xi)}{\partial \xi_i} = \frac{\partial \Upsilon_j(\Xi(\xi))}{\partial \xi_i} = \sum_{k=1}^{2N} \frac{\partial \Upsilon_j(\Xi)}{\partial \Xi_k}\frac{\partial \Xi_k(\xi)}{\partial \xi_i} \\ &= \sum_{k=1}^{2N} \left(\frac{\partial \Xi}{\partial \xi}\right)_{ik} \left(\frac{\partial \Upsilon}{\partial \Xi}\right)_{kj} = \left(\frac{\partial \Xi}{\partial \xi}\frac{\partial \Upsilon}{\partial \Xi}\right)_{ij}\ .\end{aligned} \tag{2.20}$$

This is the composition claimed for the Jacobians claimed in (3.10). It is clear that the Jacobian of a sum of two transformations is the sum of the Jacobian matrices of the individual transformations. □

**Theorem 2.6.** *Assume that $G$ does not depend explicitly on $\lambda$, and that the equation* (2.17) *has a solution that is analytic in $\lambda$. Then $\xi(\lambda)$ can be written in terms of an initial condition $\xi = \xi(0)$ at $\lambda = 0$ in the form*[2]

$$\xi_i(\lambda) = e^{-\lambda D_G}\xi_i\ . \tag{2.21}$$

*Furthermore, the Jacobian matrix $J = \frac{\partial \xi(\lambda)}{\partial \xi}$ of the transformation $\xi \mapsto \xi(\lambda)$ is $J = e^{-\lambda K}$ where $K = J_{\xi \to D_G \xi}$. The Jacobian $J$ satisfies*

$$J^{\mathrm{tr}}\,\Gamma J = \Gamma\ , \quad \textit{and} \quad \det J = 1\ . \tag{2.22}$$

*Proof.* We can rewrite the identity (2.22) for $J$ as

$$e^{-\lambda K^{\mathrm{tr}}}\Gamma e^{-\lambda K} = \Gamma\ .$$

[2]While the analyticity we assume here is sufficient, it is not necessary for the solution to be represented as (2.21). We do not discuss this technical question.

One can integrate $\dfrac{d\xi_i(\lambda)}{d\lambda}$ from an initial value at $\lambda = 0$ where $\xi(0) = \xi$. Using equation (2.17) we find

$$\xi_i(\lambda) = \xi_i + \int_0^\lambda d\lambda_1 \frac{d\xi_i(\lambda_1)}{d\lambda_1} = \xi_i - \int_0^\lambda d\lambda_1\, D_{G(\xi(\lambda_1))}\, \xi_i(\lambda_1)\ . \tag{2.23}$$

The function $C(\xi) = D_{G(\xi)}\xi_i$ is a function on phase space, so by Theorem 2.4 it satisfies the differential equation $\frac{dC(\xi(\lambda))}{d\lambda} = -D_{G(\xi(\lambda))}C(\xi(\lambda))$. Inserting this into (2.23), and integrating the first order term that is independent of $\lambda_1$, we find

$$\begin{aligned}\xi_i(\lambda) &= \xi_i - \int_0^\lambda d\lambda_1\, D_{G(\xi)}\, \xi_i + \int_0^\lambda d\lambda_1 \int_0^{\lambda_1} d\lambda_2\, D_{G(\xi(\lambda_2))} D_{G(\xi(\lambda_2))}\, \xi_i(\lambda_2) \\ &= \xi_i - \lambda D_{G(\xi)}\, \xi_i + \int_0^\lambda d\lambda_1 \int_0^{\lambda_1} d\lambda_2\, D_{G(\xi(\lambda_2))} D_{G(\xi(\lambda_2))}\, \xi_i(\lambda_2)\ .\end{aligned}$$

We can continue to iterate this process, replacing the integral of $D^k_{G(\xi(\lambda_1))}\xi(\lambda_1)$ by $D^k_{G(\xi)}\xi$, plus an integral of $D^{k+1}_{G(\xi(\lambda_1))}\xi(\lambda_1)$. Assuming the series converges, we have the result as a transformation acting on the initial conditions,

$$\xi_i(\lambda) = \sum_{k=0}^{\infty} (-1)^k \left( \int_0^\lambda \int_0^{\lambda_1} \int_0^{\lambda_2} \cdots \int_0^{\lambda_{k-1}} d\lambda_1 \cdots d\lambda_k \right) D^k_{G(\xi)} \xi_i\ . \tag{2.24}$$

Here we use the convention that the $k = 0$ term is $\xi_i$. The variables of integration are ordered in the term of degree $k$ by $0 \leqslant \lambda_k \leqslant \lambda_{k-1} \leqslant \cdots \leqslant \lambda_1 \leqslant \lambda$. Note that the integrand is a symmetric function of $\lambda_1, \ldots, \lambda_k$, so as there are $k!$ possible orderings of the variables in the $k$-dimensional cube,

$$\int_0^\lambda \int_0^{\lambda_1} \int_0^{\lambda_2} \cdots \int_0^{\lambda_{k-1}} d\lambda_1 \cdots d\lambda_k = \frac{1}{k!} \left( \int_0^\lambda d\lambda_1 \right)^k = \frac{\lambda^k}{k!}\ . \tag{2.25}$$

Hence

$$\xi_i(\lambda) = \sum_{k=0}^{\infty} \frac{(-\lambda D_G)^k}{k!} \xi_i = e^{-\lambda D_G} \xi_i\ , \tag{2.26}$$

which is (2.21).

Let us find the Jacobian $J_{\xi \to D_G \xi}$ of the transformation $\xi \to D_G \xi$ acting on phase space. Since

$$(D_G\xi)_j = D_G\, \xi_j = \sum_{\alpha,\beta} \frac{\partial G}{\partial \xi_\alpha}\, \Gamma_{\alpha\beta}\, \frac{\partial \xi_j}{\partial \xi_\beta} = \sum_\alpha \frac{\partial G}{\partial \xi_\alpha}\, \Gamma_{\alpha j}\ , \tag{2.27}$$

the Jacobian matrix $J_{\xi \to D_G}$ has entries

$$(J_{\xi \to D_G})_{ij} = \frac{\partial (D_G \xi_j)}{\partial \xi_i} = \sum_\alpha \frac{\partial^2 G}{\partial \xi_i \partial \xi_\alpha}\, \Gamma_{\alpha j} = (\mathfrak{h}\Gamma)_{ij}\ . \tag{2.28}$$

Here we introduce the matrix $\mathfrak{h}$, defined as the matrix of second derivatives of the function $G$. The matrix $\mathfrak{h}$ is also called the Hessian matrix of $G$. The entries of $\mathfrak{h}$ are

$$\boxed{\mathfrak{h}(\xi)_{ij} = \frac{\partial^2 G(\xi)}{\partial \xi_i \partial \xi_j}} . \tag{2.29}$$

Note that $\mathfrak{h}$ is real and symmetric, while $\Gamma$ is real and antisymmetric. Thus

$$(\mathfrak{h}\Gamma)^{\mathrm{tr}} = -\Gamma\mathfrak{h} . \tag{2.30}$$

If $J_T$ and $J_S$ are the Jacobians of the transformations $T$ and $S$, then the Jacobian of $\lambda T$ is $\lambda J_T$. Also the Jacobian $J_{T+S}$ is $J_T + J_S$. Hence using Proposition 2.2, the Jacobian $J$ of $e^{-\lambda D_G}$ is $e^{-\lambda\mathfrak{h}\Gamma}$. Then

$$J^{\mathrm{tr}}\,\Gamma\,J = e^{\lambda\Gamma\mathfrak{h}}\,\Gamma\,e^{-\lambda\mathfrak{h}\Gamma} \tag{2.31}$$

But $(\Gamma\mathfrak{h})^n\Gamma = \Gamma(\mathfrak{h}\Gamma)^n$, so $e^{\Gamma\mathfrak{h}}\Gamma = \Gamma e^{\mathfrak{h}\Gamma}$. As a consequence, we infer

$$J^{\mathrm{tr}}\,\Gamma J = e^{\lambda\Gamma\mathfrak{h}}\Gamma e^{-\lambda\mathfrak{h}\Gamma} = e^{\lambda\Gamma\mathfrak{h}} e^{-\lambda\Gamma\mathfrak{h}}\Gamma = \Gamma . \tag{2.32}$$

Finally, the relation (2.32) along with $\det J = \det J^{\mathrm{tr}}$ and $\det\Gamma \neq 0$ ensure that $(\det J)^2 = 1$, so that $\det J = \pm 1$. As $J$ is an $N\times N$ matrix, its determinant $\det J$ is a continuous function of $\lambda$. If $\lambda = 0$, then $J = I$ and $\det J = +1$. Hence by continuity $\det J = 1$ for all $\lambda$. □

Now we consider the case that $G(\xi(\lambda), \lambda)$ depends explicitly on $\lambda$. We see that solutions to equations of Hamiltonian type also lead to $\det J = 1$. However in this case $G$ is no longer conserved and the Jacobian is not the usual exponential. It can be written as a "ordered exponential" as defined below. The equation of evolution now is,

$$\frac{d}{d\lambda}\,\xi_i(\lambda) = -D_{G(\xi(\lambda),\lambda)}\,\xi_i(\lambda) . \tag{2.33}$$

**Theorem 2.7.** *Assume that the equation* (2.17) *has a solution* $\xi(\lambda)$ *that is analytic in* $\lambda$. *Then the Jacobian matrix* $J = \frac{\partial\xi(\lambda)}{\partial\xi}$ *of the transformation* $\xi \to \xi(\lambda)$ *satisfies*

$$J^{\mathrm{tr}}\,\Gamma J = \Gamma , \quad \textit{and} \quad \det J = 1 . \tag{2.34}$$

*Proof.* One can integrate $\frac{d\xi_i(\lambda)}{d\lambda}$ from an initial value of $\lambda$ (say $\lambda = 0$) at which $\xi(0) = \xi$ to $\lambda$.

$$\xi_i(\lambda) = \xi_i + \int_0^\lambda d\lambda_1 \frac{d\xi_i(\lambda_1)}{d\lambda_1} = \xi_i - \int_0^\lambda d\lambda_1\, D_{G(\xi(\lambda_1),\lambda_1)}\xi_i(\lambda_1) . \tag{2.35}$$

Repeating this expansion for the first derivative, we find

$$\xi_i(\lambda) = \xi_i - \int_0^\lambda d\lambda_1\, D_{G(\xi(\lambda_1),\lambda_1)}\,\xi_i + \int_0^\lambda\int_0^{\lambda_1} d\lambda_1 d\lambda_2\, D_{G(\xi(\lambda_1),\lambda_1)} D_{G(\xi(\lambda_2),\lambda_2)}\xi_i(\lambda_2) . \tag{2.36}$$

And continuing to iterate this procedure we obtain

$$\xi_i(\lambda) = \sum_{k=0}^{\infty}(-1)^k \int_0^\lambda\int_0^{\lambda_1}\int_0^{\lambda_2}\cdots\int_0^{\lambda_{k-1}} d\lambda_1\cdots d\lambda_k\, D_{G(\xi(\lambda_1),\lambda_1)} D_{G(\xi(\lambda_2),\lambda_2)}\cdots D_{G(\xi(\lambda_k),\lambda_k)}\,\xi_i . \tag{2.37}$$

The difference between this formula and the case with $G(\xi(\lambda))$ only depending on $\lambda$ through the dependence of $\xi(\lambda)$, is the fact that we can no longer separate the integration over $\lambda_1, \ldots, \lambda_k$ from the action of $D_{G(\xi(\lambda_1),\lambda_1)}$, etc. Thus it is not apparent that one can sum the terms into an exponential. Freeman Dyson suggested a way solve this problem, by introducint to write this a "$\lambda$-ordered product." This is defined to be symmetric in the arguments, so the product is commutative. It is also compatible with the case when $G$ does not depend explicitly on $\lambda$. Let

$$\left(D_{G(\xi(\lambda_1),\lambda_1)}\, D_{G(\xi(\lambda_2),\lambda_2)}\right)_- = \begin{cases} D_{G(\xi(\lambda_1),\lambda_1)}\, D_{G(\xi(\lambda_2),\lambda_2)}\ , & \text{if } \lambda_1 \geqslant \lambda_2 \\ D_{G(\xi(\lambda_2),\lambda_1)}\, D_{G(\xi(\lambda_2),\lambda_2)}\ , & \text{if } \lambda_2 \geqslant \lambda_1 \end{cases} \tag{2.38}$$

The $-$ denotes the decrease of $\lambda$ from left to right. One then defines the $k$-fold ordered product to have the order of decreasing $\lambda$'s from left to right, out of the $k!$ possible orders. Thus we can express the sum as the $\lambda$-ordered exponential,

$$\xi_i(\lambda) = \left(e^{-\int_0^\lambda D_{G(\xi(\lambda_1),\lambda_1)}\, d\lambda_1}\right)_- \xi_i\ . \tag{2.39}$$

This expression comes up a great deal in quantum theroy, where it is known as "Dyson's formula."

In order to see that this form of $\xi \to \xi(\lambda)$ is canonical, we need to modify the previous argument to compute the Jacobian. One can follow the previous argument, but in the case that $G$ depends explicitly on $\lambda$, the Hessian matrix $\mathfrak{h}(\xi(\lambda), \lambda)$, defined in (2.29), depends explicitly on $\lambda$. Following the previous argument in the proof of Theorem 2.6 , we find that now

$$J = \left(e^{-\int_0^\lambda \mathfrak{h}(\lambda_1),\lambda_1\, \Gamma\, d\lambda_1}\right)_-\ . \tag{2.40}$$

As before,

$$J^{\mathrm{tr}}\Gamma\, J = \Gamma\ . \tag{2.41}$$

As in the proof of Theorem 2.6, we establish $\det J = 1$. □

## 3. Canonical Transformations

In this section we consider invertible changes of coordinates

$$\xi \mapsto \Xi = \Xi(\xi)\ , \tag{3.1}$$

where $\xi$ and $\Xi$ are $2N$-dimensional vectors. Denote the components of $\xi$ by $\xi_i$, and the corresponding components of $\Xi$ by $\Xi_i$.

This transformation $\xi \mapsto \Xi$ is not necessarily linear. In the study of the simple harmonic oscillator, we considered the solution of Hamilton's equations giving such a change of coordinates, from the initial condition $\xi$ to the position $\Xi$ at some other time $t$. In that example the transformation $\xi \mapsto \Xi$ was linear, and given by a matrix $e^{tT}$ acting on the vector $\xi$ to give the transformation $\Xi = e^{tT}\xi$. In that case we found that the Jacobian $J$ of the transformation satisfied $J^{\mathrm{tr}}\,\Gamma J = \Gamma$, and that $\det J = 1$. Now we consider more general transformations which in general are non-linear. However we see that the same properties are very natural. This leads us to the definition of a canonical transformation.

### 3.1. **What is a Canonical Transformation?**

**Definition 3.1.** *The transformation $\xi \mapsto \Xi(\xi)$ on phase space is canonical, if*

$$[\Xi_i\,,\,\Xi_j]_\xi = \Gamma_{ij}\;. \tag{3.2}$$

In other words, a canonical transformation preserves the fundamental Poisson brackets.

**Proposition 3.2.** *A transformation $\xi \mapsto \Xi(\xi)$ on phase space is canonical, if and only if its Jacobian $J = \frac{\partial \Xi}{\partial \xi}$ satisfies*

$$\boxed{J^{\mathrm{tr}}\,\Gamma J = \Gamma}\,, \quad \textit{and} \quad \boxed{\det J = \pm 1}\;. \tag{3.3}$$

**Remark 3.3.** A canonical transformation is said to be *proper* if $\det J = +1$. We will see that most of the canonical transformations that we study in these notes come from families of transformations that have the property $\det J = +1$. In fact we infer from Theorem 2.6 every solution of an equation of Hamiltonian type defines a transformation $\xi \to \xi(\lambda)$ that is a proper, canonical transformation.

**Remark 3.4.** We could have defined a canonical transformation as a transformation $\xi \mapsto \Xi$ that has a Jacobian $J$ that satisfies $J^{\mathrm{tr}}\Gamma J = \Gamma$. Then the proposition would be to show that such transformation preserves the fundamental Poisson brackets (3.1).

*Proof.* It is straightforward to compute the Poisson brackets in question in terms of the Jacobian of the transformation. Namely,

$$[\Xi_i\,,\,\Xi_j]_\xi = \sum_{k,\ell=1}^{2N} \frac{\partial \Xi_i}{\partial \xi_k}\Gamma_{k\ell}\frac{\partial \Xi_j}{\partial \xi_\ell} = \sum_{k,\ell=1}^{2N} J_{ki}\Gamma_{k\ell}J_{\ell j} = \left(J^{\mathrm{tr}}\Gamma J\right)_{ij}\;. \tag{3.4}$$

Thus the condition $J^{\mathrm{tr}}\Gamma J = \Gamma$ is equivalent to the statement that $[\Xi_i\,,\,\Xi_j]_\xi = \Gamma_{ij}$. Note that the matrix $\Gamma$ has determinant $(-1)^N \neq 0$, and $\det J^{\mathrm{tr}} = \det J$. By taking the determinant of (3.4), we infer that $\det(J^{\mathrm{tr}}\Gamma J) = (\det J)^2 \det\Gamma = \det\Gamma$. Hence factoring out $\det\Gamma$, we have $\det J = \pm 1$. ☐

We now show that a canonical transformation $\xi \mapsto \Xi$ preserves the Poisson bracket of any two functions $A(\Xi(\xi))$ and $B(\Xi(\xi))$ on phase space. What does this mean? We can compute the Poisson bracket in two ways. One way is to find $[A(\Xi(\xi))\,,\,B(\Xi(\xi))]_\xi$ in the old coordinates. Alternatively we can compute $[A(\Xi)\,,\,B(\Xi)]_\Xi$ in the new coordinates. We show that the two answers agree.

**Proposition 3.5** (**Canonical Transformations Preserve All Poisson Brackets).** *Let $\xi \mapsto \Xi$ be a canonical transformation on $\mathbb{R}^{2N}$, and let $A, B$ be differentiable functions on $\mathbb{R}^{2N}$. Then*

$$\boxed{[A(\Xi(\xi))\,,\,B(\Xi(\xi))]_\xi = [A(\Xi)\,,\,B(\Xi)]_\Xi}\;. \tag{3.5}$$

*Proof.* Again this is an elementary computation, which only uses the chain rule for differentiation. We have

$$\begin{aligned}
[A(\Xi(\xi))\,,\,B(\Xi(\xi))]_\xi &= \sum_{i,j,k,l=1}^{2N} \frac{\partial A(\Xi(\xi))}{\partial \Xi_i}\,\frac{\partial \Xi(\xi)_i}{\partial \xi_j}\,\Gamma_{jk}\,\frac{\partial B(\Xi(\xi))}{\partial \Xi_l}\,\frac{\partial \Xi(\xi)_l}{\partial \xi_k} \\
&= \sum_{i,l=1}^{2N} \frac{\partial A(\Xi(\xi))}{\partial \Xi_i}\,\Gamma_{il}\,\frac{\partial B(\Xi(\xi))}{\partial \Xi_l} = [A(\Xi)\,,\,B(\Xi)]_\Xi\ . \qquad (3.6)
\end{aligned}$$

In the second equality we use the identity $J^{\mathrm{tr}}\Gamma J = \Gamma$ given in (3.3) for the Jacobian of a canonical transformation. The third equality is the definition of the Poisson bracket in the coordinates $\Xi$. □

**Proposition 3.6.** *If $J$ is the Jacobian of a canonical transformation, then $J^{\mathrm{tr}}$ and $J^{-1}$ are also Jacobians of canonical transformations. The product of two canonical transformations is canonical.*

*Proof.* Assume $J^{\mathrm{tr}}\Gamma J = \Gamma$, and multiply this equation on the left by $J\Gamma$ and on the right by $J^{-1}\Gamma^{-1}$ to give $(J\Gamma)(J^{\mathrm{tr}}\Gamma J)(J^{-1}\Gamma^{-1}) = (J\Gamma)\Gamma(J^{-1}\Gamma^{-1})$. This simplifies, using $\Gamma^2 = -I$, to $J\Gamma J^{\mathrm{tr}} = -\Gamma^{-1} = \Gamma$. Thus $J^{\mathrm{tr}}$ is the Jacobian of a canonical transformation.

In order to say the same for $J^{-1}$, take the inverse of $J\Gamma J^{\mathrm{tr}} = \Gamma$ to obtain

$$(J^{\mathrm{tr}})^{-1}\Gamma J^{-1} = \Gamma\ .$$

We claim that $(J^{\mathrm{tr}})^{-1} = (J^{-1})^{\mathrm{tr}}$. In fact,

$$\left(J^{\mathrm{tr}}\right)^{-1} = \left(JJ^{-1}\right)^{\mathrm{tr}}\left(J^{\mathrm{tr}}\right)^{-1} = \left(J^{-1}\right)^{\mathrm{tr}} J^{\mathrm{tr}}\left(J^{\mathrm{tr}}\right)^{-1} = \left(J^{-1}\right)^{\mathrm{tr}}\ .$$

So $J^{-1}$ is the Jacobian of a canonical transformation. Finally, assume that $J_1, J_2$ are Jacobians of canonical transformations. Then

$$(J_1J_2)^{\mathrm{tr}}\Gamma J_1J_2 = J_2^{\mathrm{tr}}(J_1^{\mathrm{tr}}\Gamma J_1)J_2 = J_2^{\mathrm{tr}}\Gamma J_2 = \Gamma\ ,$$

as claimed. □

3.2. **Groups.** A *group* $\mathfrak{G}$ is a collection of objects with the following properties:

- **Associative Multiplication:** If $g_1, g_2 \in \mathfrak{G}$, then there is a product $g_1g_2 \in \mathfrak{G}$ that satisfies $(g_1g_2)g_3 = g_1(g_2g_3)$.
- **Identity:** There is a unit $e \in \mathfrak{G}$, such that for all $g \in \mathfrak{G}$, one has $ge = eg = g$.
- **Inverse:** If $g \in \mathfrak{G}$, then there exists $g^{-1} \in \mathfrak{G}$ such that $gg^{-1} = g^{-1}g = e$.

Often one can parameterize an abelian subgroup of $\mathfrak{G}$ by a real parameter $\lambda$, so that $g(\lambda_1)g(\lambda_2) = g(\lambda_1 + \lambda_2)$. In this case $g(\lambda)^{-1} = g(-\lambda)$, and $g(0) = e$. More generally the space $\mathbb{R}^N$ is an elementary example of the translation group on $\mathbb{R}^N$, with the group multiplication law being given by addition of vectors.

Another class of groups arises from non-singular matrices, with the group law given by matrix multiplication. A square matrix is invertible if and only if its determinant is non-vanishing. The set of all $N \times N$ matrices with non-zero determinant is called the *general-linear* group of $N \times N$ matrices. One denotes this group $GL(N, \mathbb{C})$, where $\mathbb{C}$ denotes complex matrix entries. Let us list some very important subgroups of $GL(N, \mathbb{C})$, all of which enter physics:

(1) The *special linear group* $SL(N, \mathbb{C})$ are matrices in $GL(N, \mathbb{C})$ with determinant $+1$.
(2) The *modular group* $SL(N, \mathbb{Z})$ is the subgroup of $SL(N, \mathbb{C})$ with integer matrix elements.
(3) The *unitary group* $U(N)$ consists of matrices in $GL(N, \mathbb{C})$ that are unitary.
(4) The *special unitary group* $SU(N)$ consists of unitary matrices with determinant $+1$.
(5) The *special orthogonal group* $SO(N, \mathbb{C})$ consists of orthogonal matrices with determinant $+1$.
(6) The *real orthogonal group* $SO(N, \mathbb{R})$, or simply $SO(N)$ consists of real orthogonal matrices with determinant $+1$.

3.2.1. *Examples.* The simplest examples of groups arise from sets of $N \times N$ matrices. And a special class of groups are those made up from unitary matrices $U$, namely matrices for which $U^*U = I$. A matrix $S$ is skew-hermitian if $S^* = -S$, where $S^*$ denotes the hermitian adjoint of $S$. A matrix $S$ is skew-hermitian if and only if the matrix $H = iS$ is hermitian, where $i = \sqrt{-1}$. Every unitary matrix is the exponential of a skew-hermitian matrix, $U = e^S$. In particular for a real parameter $\lambda$, the matrices

$$U(\lambda) = e^{-\lambda S} = e^{i\lambda H} = \sum_{n=0}^{\infty} \frac{(-\lambda)^n}{n!} S^n \ ,$$

are a one-parameter, additive, unitary group. As $U(\lambda)$ is normal, it has an orthonormal basis of eigenvectors. The eigenvalues of $U(\lambda)$ have absolute value 1, so they have the form $\mu_j = e^{i\lambda_j(H)}$ where $\lambda_j(H)$ is an eigenvalue of $H = H^*$.

3.2.2. *Lie Groups.* In case there are a number of generators $G_1, \ldots, G_k$, there may be a set of constants $f_{ijk}$, such that the Poisson bracket relations for all the pairs of the $G_j$'s can be written as

$$[G_i \, , \, G_j]_\xi = \sum_k f_{ijk} G_k \ . \tag{3.7}$$

One says that the Poisson bracket relations *close.* The effect of this is that one can express the various products defined by the Poisson bracket in terms of a linear combination of the $G_j$'s. In other words, the $G_i$'s generate the algebra. The constants $f_{ijk}$ are called the *structure constants* of the algebra. The other fundamental property of a Lie algebra is the Jacobi identity which says that the derivative $D_A$, defined in (2.3), acts as a derivative, with respect to the product rule given by the Poisson bracket,

$$D_A \, [B \, , \, C]_\xi = [D_A B \, , \, C]_\xi + [B \, , \, D_A C]_\xi \ . \tag{3.8}$$

We come back in §6 to show that this Jacobi identity holds for the Poisson bracket.

3.3. **Canonical Transformations Form a Group.** Transformations on phase space have a natural multiplication, given by the *composition* of the two transformations. If $\Xi$ and $\Upsilon$ are the transformations

$$\xi \mapsto \Xi(\xi) , \quad \text{and} \quad \Xi \mapsto \Upsilon(\Xi) ,$$

then the composition product $\Upsilon \circ \Xi$ is defined as the transformation

$$\xi \mapsto \Upsilon \circ \Xi , \quad \text{namely} \quad (\Upsilon \circ \Xi)(\xi) = \Upsilon(\Xi(\xi)) . \tag{3.9}$$

**Proposition 3.7.** *Canonical transformation with the product in* (3.9) *form a group.*

*Proof.* Composition of transformations is associative, so the composition product qualifies a group product. The Jacobian of a canonical transformation has determinant equal to $\pm 1$, so by the implicit function theorem each such function is invertible, giving the inverse canonical transformation. In fact the Jacobian of the inverse is the inverse of the Jacobian. Lastly the identity transformation has Jacobian equal to the unit matrix, so (3.1) is satisfied trivially and the identity is a canonical transformation. Thus we have all the necessary properties for the set of canonical transformations to be a group.

By Proposition 3.7, the Jacobian matrix of the composition of transformations if the product of Jacobians. This product is associative and invertible, since the Jacobians of canonical transformations have determinant $\pm 1$. □

## 4. A Few Examples of Canonical Transformations

We now give some specific examples of canonical transformations. In later sections we generalize these examples and give many others.

4.1. **Translation.** Clearly the transformation

$$q_j \mapsto Q_j = q_j + \lambda_j , \quad p_j \mapsto P_j = p_j + \mu_j , \tag{4.1}$$

is canonical for any choice of $\lambda_j$ and $\mu_j$. In fact $J = I$, so (3.3) is clearly satisfied.

4.2. **Linear Transformations that do not mix $q$ with $p$.** Here we combine together a number of examples, and investigate a family of invertible linear transformations that do not mix $q$'s with $p$'s. Consider transformation defined by invertible matrices $\Lambda$ and $\Lambda'$ that on $N$ coordinates $q$ and $N$ coordinates $p$ as,

$$Q = \Lambda q , \quad \text{and} \quad P = \Lambda' p . \tag{4.2}$$

The corresponding transformation $\xi \mapsto \Xi$ on phase space is

$$\boxed{\Xi = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda' \end{pmatrix} \xi} . \tag{4.3}$$

We claim that for this linear transformation, $J^{\mathrm{tr}} = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda' \end{pmatrix}$; this is the case as

$$\left(J^{\mathrm{tr}}\right)_{ij} = J_{ji} = \frac{\partial \Xi_i}{\partial \xi_j} = \frac{\partial \left( \left( \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda' \end{pmatrix} \xi \right)_i \right)}{\partial \xi_j} = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda' \end{pmatrix}_{ij} \ . \tag{4.4}$$

**Proposition 4.1.** *The transformation* (4.3) *is canonical if and only if* $\Lambda' = \left(\Lambda^{\mathrm{tr}}\right)^{-1}$, *in which case*

$$J^{\mathrm{tr}} = \begin{pmatrix} \Lambda & 0 \\ 0 & \left(\Lambda^{\mathrm{tr}}\right)^{-1} \end{pmatrix} \ . \tag{4.5}$$

*Proof.* Note that

$$\Lambda' = \left(\Lambda^{-1}\right)^{\mathrm{tr}} = \left(\Lambda^{\mathrm{tr}}\right)^{-1} \ . \tag{4.6}$$

Therefore the condition that $J$ is the Jacobian of a canonical transformation is

$$J^{\mathrm{tr}} \Gamma J = \Gamma \ , \quad \text{or} \quad \begin{pmatrix} 0 & \Lambda \Lambda'^{\,\mathrm{tr}} \\ -\Lambda' \Lambda^{\mathrm{tr}} & 0 \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \ . \tag{4.7}$$

Therefore we conclude that for this class of canonical transformations (4.5) holds. □

**Linear Example 1. Scaling:** Suppose $\Lambda = M = M^{\mathrm{tr}} > 0$ is a real, symmetric, strictly-positive matrix. Then the transformation in (4.3) is given the special name of a *scaling*, and

$$\boxed{J^{\mathrm{tr}} = J = \begin{pmatrix} M & 0 \\ 0 & M^{-1} \end{pmatrix} \ , \qquad \textbf{(Scaling Transformation).}} \tag{4.8}$$

For a scaling, the configuration space coordinates $q$ and the momenta $p$ have inverse transformation laws. The scaling can affect an individual coordinate or every coordinate. The classical scaling corresponds to a diagonal matrix $M$, but other more general scalings are possible as well.

**Linear Example 2. Rotation:** Let us assume that $\Lambda$ is a real, invertible matrix satisfying $\Lambda^{-1} = \Lambda^{\mathrm{tr}}$. In other words, $\Lambda$ is a real orthogonal matrix, which we denote $R^{\mathrm{tr}}$, with $R$ suggesting "rotation." In this circumstance,

$$\boxed{J^{\mathrm{tr}} = J^{-1} = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix} \ , \qquad \textbf{(Rotation).}} \tag{4.9}$$

For a rotation, both $q$ and $p$ rotate by the same transformation.

**Remark: Every real, invertible matrix $\Lambda$ can be written $\boldsymbol{\Lambda = MR}$.** Here $M > 0$ is a symmetric, positive matrix and $R$ is a real orthogonal matrix. The representation $\Lambda = MR$ is the polar decomposition of the $\Lambda$. As $\Lambda^{\mathrm{tr}} = \Lambda^*$, define $M$ as the positive square root $M = (\Lambda^{\mathrm{tr}}\Lambda)^{1/2}$, and $R = M^{-1}\Lambda$. Thus every real, invertible $N \times N$ matrix $\Lambda$ gives a $2N \times 2N$ Jabobian that does not mix $q$ with $p$; this Jacobian is the product of a Jacobian of a rotation, followed by a Jacobian of a scaling,

$$J^{\mathrm{tr}} = \begin{pmatrix} \Lambda & 0 \\ 0 & (\Lambda^{\mathrm{tr}})^{-1} \end{pmatrix} = \begin{pmatrix} M & 0 \\ 0 & M^{-1} \end{pmatrix} \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix} . \tag{4.10}$$

4.3. **The Simple Harmonic Oscillator Mixes $\boldsymbol{q}$ with $\boldsymbol{p}$.** Let $\Omega$ be an $N \times N$ real, symmetric matrix with positive eigenvalues, and consider the $2N \times 2N$ matrix,

$$T = \begin{pmatrix} 0 & I \\ -\Omega^2 & 0 \end{pmatrix} . \tag{4.11}$$

Any two functions $f(\Omega)$ and $g(\Omega)$ commute. So

$$e^{tT} = \begin{pmatrix} \cos(\Omega t) & \Omega^{-1}\sin(\Omega t) \\ -\Omega\sin(\Omega t) & \cos(\Omega t) \end{pmatrix} , \quad \text{and} \quad e^{tT^{\mathrm{tr}}} = \begin{pmatrix} \cos(\Omega t) & -\Omega\sin(\Omega t) \\ \Omega^{-1}\sin(\Omega t) & \cos(\Omega t) \end{pmatrix} . \tag{4.12}$$

Similarly,

$$\begin{aligned} e^{tT}\Gamma e^{tT^{\mathrm{tr}}} &= \begin{pmatrix} \cos(\Omega t) & \Omega^{-1}\sin(\Omega t) \\ -\Omega\sin(\Omega t) & \cos(\Omega t) \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} \cos(\Omega t) & -\Omega\sin(\Omega t) \\ \Omega^{-1}\sin(\Omega t) & \cos(\Omega t) \end{pmatrix} \\ &= \begin{pmatrix} -\Omega^{-1}\sin(\Omega t) & \cos(\Omega t) \\ -\cos(\Omega t) & -\Omega\sin(\Omega t) \end{pmatrix} \begin{pmatrix} \cos(\Omega t) & -\Omega\sin(\Omega t) \\ \Omega^{-1}\sin(\Omega t) & \cos(\Omega t) \end{pmatrix} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \\ &= \Gamma . \end{aligned} \tag{4.13}$$

It follows that the family of transformations (parameterized by $t$) of the form

$$\Xi = e^{tT}\xi = J^{\mathrm{tr}}\xi \tag{4.14}$$

are a family of canonical transformations.

But $\Xi = e^{tT}\xi = J^{\mathrm{tr}}\xi$ is just the solution of Hamilton's equations for the simple harmonic oscillator described in Equations (52) and (54) of §II.1.4 in the notes on Normal Modes. One also recovers the general oscillator solution by conjugating this transformation with the scaling (4.8) (with $M$ replaced by $M^{1/2}$), as described in Equation (53) of the notes on Normal Modes. Then we obtain $\Xi = e^{tT}\xi = J^{\mathrm{tr}}\xi$ with the new Jacobian having the transpose

$$\begin{aligned} J^{\mathrm{tr}} &= \begin{pmatrix} M^{-1/2} & 0 \\ 0 & M^{1/2} \end{pmatrix} \begin{pmatrix} \cos(\Omega t) & \Omega^{-1}\sin(\Omega t) \\ -\Omega\sin(\Omega t) & \cos(\Omega t) \end{pmatrix} \begin{pmatrix} M^{1/2} & 0 \\ 0 & M^{-1/2} \end{pmatrix} \\ &= \begin{pmatrix} M^{-1/2}\cos(\Omega t)M^{1/2} & M^{-1/2}\Omega^{-1}\sin(\Omega t)M^{-1/2} \\ -M^{1/2}\Omega\sin(\Omega t)M^{1/2} & M^{1/2}\cos(\Omega t)M^{-1/2} \end{pmatrix} . \end{aligned} \tag{4.15}$$

4.4. **Some Discrete Canonical Transformations.** Consider the transformation

$$\boxed{Q = p\ , \quad P = -q\ , \quad \text{or} \quad \Xi = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \xi = \Gamma\xi}\ . \tag{4.16}$$

As $J^{\mathrm{tr}} = \Gamma$, one has $J\,\Gamma J^{\mathrm{tr}} = -\Gamma^3 = \Gamma$. Hence this transformation is canonical. In fact, this transformation is a special case of the canonical, oscillator evolution transformation (4.15), corresponding to the choice: $M = \Omega = I$ and $t = \frac{\pi}{2}$.

An amusing property of the transformation (4.16) is that one can consider this transformation as a classical-mechanics analog of the Fourier transform: the position coordinate $q$ is transformed to the dual momentum coordinate $p$, while $p$ is transformed to $-q$. Like Fourier transformation, this transformation has period 4; this corresponds to $\Gamma^4 = I$. As the canonical transformations are a group, the Jacobian $J^{\mathrm{tr}}$ corresponding to the 4-fold application of (4.16) is $J^{\mathrm{tr}} = \Gamma^4 = I$.

On the other hand, if one applies the transformation twice, then one obtains the canonical transformation with the squared Jacobian matrix, namely the discrete reflection transformation,

$$\boxed{Q = -q\ , \quad P = -p\ , \qquad \text{namely} \quad \Xi = \begin{pmatrix} -I & 0 \\ 0 & -I \end{pmatrix} \xi = -\xi}\ . \tag{4.17}$$

In this case $J^{\mathrm{tr}} = \Gamma^2 = -I$

As another variant, consider the canonical, oscillator evolution transformation (4.15), corresponding to the choice: $M = \Omega = I$ and $t = \frac{3\pi}{2}$. This yields the canonical transformation with $J^{\mathrm{tr}} = -\Gamma$, or

$$\boxed{Q = -p\ , \quad P = q\ , \qquad \text{namely} \quad \Xi = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix} \xi = -\Gamma\xi}\ . \tag{4.18}$$

## 5. All the Examples in §4 Come from Hamilton's Equations

In this section, we analyze the solutions to various Hamilton's equations that arise from assuming that the Hamiltonians $H(\xi)$ is a homogeneous quadratic function of $\xi$. In this case the Hamilton equations are linear. We solve them in closed form, as we did for the simple harmonic oscillator. We recover all the examples of canonical transformations in §4, as well as some others.

For the oscillator we found that the original Hamilton's equation

$$\dot{\xi}(t) = -D_H \xi(t)\ , \tag{5.1}$$

reduces to the linear, first-order differential equation

$$\dot{\xi}(t) = T\xi(t)\ , \quad \text{where } T \text{ is a fixed matrix,} \tag{5.2}$$

whose entries are independent of $t$ and of $\xi$. The solution $\xi(t) = e^{tT}\xi(0)$ defines a transformation $\xi \mapsto \Xi$, where $\xi = \xi(0)$ and $\Xi = \xi(t)$. Moreover, this is a canonical transformation.

The Jacobian of the transformation $\xi \mapsto \Xi$ is $e^{tT^{\mathrm{tr}}}$, so the condition $J^{\mathrm{tr}}\Gamma J = \Gamma$ that the transformation is canonical can be written

$$e^{tT}\Gamma e^{tT^{\mathrm{tr}}} = \Gamma\ . \tag{5.3}$$

In the case of the simple harmonic oscillator, we showed this to be the case for the solution that we found in closed form.

Here is an alternative, way to check whether the relation (5.3) holds, by reducing the condition for an exponential of $T$ to an elementary condition on $T$ itself—which is easier to verify, for we do not need to know the form of the exponential. Clearly (5.3) holds for $t = 0$. In order for this to hold also for all $t \neq 0$, it is sufficient (and also necessary) that the time derivative vanishes, namely

$$\frac{d}{dt}\left(e^{tT}\Gamma e^{tT^{\mathrm{tr}}}\right) = e^{tT}\left(T\Gamma + \Gamma T^{\mathrm{tr}}\right) e^{tT^{\mathrm{tr}}} = 0\ . \tag{5.4}$$

Evaluating the derivative at $t = 0$, we have find that we must have $T\Gamma = -\Gamma T^{\mathrm{tr}}$, or

$$\boxed{\Gamma T \Gamma = T^{\mathrm{tr}}}\ . \tag{5.5}$$

But this condition also ensures the vanishing of the derivative for all times $t$. So the relation (5.5) is equivalent to ether the relation (5.3) for $J = e^{tT^{\mathrm{tr}}}$ being the Jacobian of a canonical transformation. We call this an "infinitesimal condition" for the matrix $T$ to give a canonical transformation.

In the oscillator example,

$$T = \begin{pmatrix} 0 & M^{-1} \\ -K & 0 \end{pmatrix}\ , \quad \text{with} \quad M = M^{\mathrm{tr}} \text{ and } K = K^{\mathrm{tr}}\ , \tag{5.6}$$

and we easily check that

$$\Gamma T \Gamma = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}\begin{pmatrix} 0 & M^{-1} \\ -K & 0 \end{pmatrix}\begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \begin{pmatrix} 0 & -K \\ M^{-1} & 0 \end{pmatrix} = T^{\mathrm{tr}}\ . \tag{5.7}$$

We can generalize this elementary idea. Suppose we start from any Hamiltonian that is a real, quadratic polynomial in $\xi$. We can always write such a Hamiltonian in the following form

$$H = \frac{1}{2}\langle \xi, \mathfrak{h}\xi\rangle + \langle \alpha, \xi\rangle\ , \quad \text{where} \quad \mathfrak{h} = \mathfrak{h}^{\mathrm{tr}} = \begin{pmatrix} K & A^{\mathrm{tr}} \\ A & B \end{pmatrix}\ . \tag{5.8}$$

It is no loss of generality to take the $2N \times 2N$ matrix $\mathfrak{h}$ to be symmetric, as $\langle \xi, \mathfrak{h}\xi\rangle = \sum_{i,j} \mathfrak{h}_{ij}\xi_i\xi_j$, to which any skew part of $\mathfrak{h}$ gives 0. We write this matrix in block form. Here $\alpha$ is a $2N$-vector. We empahsize that $A, B, K$ are real, $N \times N$ matrices, and both $K = K^{\mathrm{tr}}$ and $B = B^{\mathrm{tr}}$ are symmetric. Also the vector $\alpha$ is real. We have also assumed that $H(0) = 0$. (Adding a constant to $H$ does not change Hamilton's equations.)

5.1. **The General Quadratic Hamiltonian.** In the general case, Hamiltonian's equations $\dot{\xi}(t) = -D_H \xi(t)$ simplify to become

$$\dot{\xi}(t) = \widetilde{T}\xi(t) + \beta \ , \quad \text{where} \quad \widetilde{T} = \Gamma\mathfrak{h} = \begin{pmatrix} A & B \\ -K & -A^{\mathrm{tr}} \end{pmatrix} \text{ and } \beta = \Gamma\alpha \ . \tag{5.9}$$

(The previous example (5.6) corresponds to taking $A = 0$, $B = M^{-1}$,and $\alpha = 0$, in which case $\widetilde{T} = T$ of that example.)

This is the general first-order differential equation for $\xi(t)$ with the coefficient $\widetilde{T}$ a constant $(2N \times 2N)$-matrix and the coefficient $\beta$ a constant $2N$-vector. This equation has a unique solution that agrees with the initial condition, namely

$$\boxed{\xi(t) = e^{t\widetilde{T}}\xi(0) + \int_0^t e^{(t-s)\widetilde{T}} ds\, \beta} \ . \tag{5.10}$$

Let us define $\xi = \xi(0)$ and $\Xi = e^{t\widetilde{T}}\xi + \int_0^t e^{(t-s)\widetilde{T}} ds\, \beta$. We ask under what conditions is the transformation

$$\xi \mapsto \Xi \tag{5.11}$$

canonical? The Jacobian of this transformation is $J = e^{t\widetilde{T}^{\mathrm{tr}}}$. So one finds by the same argument that led above to (5.5), that (5.11) defines a family canonical transformation from $\xi$ to $\Xi$ parameterized by $t$ if and only if $\Gamma\widetilde{T}\Gamma = \widetilde{T}^{\mathrm{tr}}$. But we compute

$$\Gamma\widetilde{T}\Gamma = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \begin{pmatrix} A & B \\ -K & -A^{\mathrm{tr}} \end{pmatrix} \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} = \begin{pmatrix} A^{\mathrm{tr}} & -K \\ M^{-1} & -A \end{pmatrix} = \widetilde{T}^{\mathrm{tr}} \ .$$

Therefore *every* $H$ of the form (5.8) gives a canonical transformation of the form (5.11). We say that the canonical transformations of the form $\xi \mapsto \Xi = e^{-\lambda D_G}\xi$ are *generated by the function* $G$, namely by a given function on phase space. We now use the parameter $\lambda$, instead of $t$, because the parameter may no longer have the interpretation of time. It may be a length, an angle, or something else. Also we use the notation $G$ for the generator, as it may not have the physical interpretation of a "Hamiltonian" or energy function. It may be a coordinate, an angular momentum, or something else.

We denote the family of transformations $T_\lambda = e^{-\lambda D_G}$ that define the solution by the symbol $T_\lambda$. This family of transformations on phase space is actually a one-parameter group for different values of $\lambda$. It has the properties:

$$T_\lambda T_\mu = T_{\lambda+\mu} \ , \quad T_0 = I \ , \quad T_\lambda^{-1} = T_{-\lambda} \ . \tag{5.12}$$

5.2. **Translation (Affine Transformation).** We claim that the canonial transformations considered in §4.1,

$$q_j \mapsto Q_j = q_j + \lambda_j \ , \quad p_j \mapsto P_j = p_j + \mu_j \ , \quad \text{for} \quad j = 1, \ldots, N \ , \tag{5.13}$$

are generated by

$$G = \sum_{j=1}^{N} (\lambda_j p_j - \mu_j q_j) \ , \quad \text{which gives} \quad -D_G = \sum_{j=1}^{N} \left( \lambda_j \frac{\partial}{\partial q_j} + \mu_j \frac{\partial}{\partial p_j} \right) \ . \tag{5.14}$$

One can also write

$$G = \langle \lambda, p \rangle - \langle \mu, q \rangle \ , \quad \text{for which} \quad -D_G = \langle \lambda, \nabla_q \rangle + \langle \mu, \nabla_p \rangle \ . \tag{5.15}$$

The exponential power series for

$$\Xi_j = e^{-D_G} \xi_j = \sum_{n=0}^{\infty} \frac{(-D_G)^n}{n!} \, \xi_j \ , \tag{5.16}$$

yields

$$Q_j = e^{-D_G} q_j = q_j - D_G q_j = q_j + \lambda_j \ , \tag{5.17}$$

and

$$P_j = e^{-D_G} p_j = p_j - D_G p_j = p_j + \mu_j \ . \tag{5.18}$$

In fact both of these exponential power series terminates after just two terms. This is the general case of a linear generating function $G$ as a function of the coordinates. One can say that a linear function $G$ generates an affine transformation on phase space.

5.3. **Generators $G$ for These Linear Canonical Transformations.** Now we find a function $G(\xi)$ on phase space that generates these transformations. Suppose that the $N \times N$ matrix $\Lambda$ can be written as $\Lambda = e^T$ for an $N \times N$ matrix $T$. Then we claim that the quadratic function on phase space equal to

$$G(\xi) = \langle p, Tq \rangle = \sum_{i,j=1}^{N} p_i T_{ij} q_j \tag{5.19}$$

generates the linear transformations above, in the sense that

$$\Xi = J\xi = e^{-D_G} \xi \ . \tag{5.20}$$

**The case $\Lambda$ is real, symmetric, and strictly positive:** In this case we find a matrix $T$ such that $\Lambda = e^T$. We call $T = \ln \Lambda$ the branch of the logarithm of $\Lambda$ with real eigenvalues. One finds this $T$ as follows: diagonalize the hermitian matrix $\Lambda$ to obtain $D = U^* \Lambda U$. As the eigenvalues of $\Lambda$ (and those of $D$) are positive by assumption, take the real logarithm of each eigenvalue to define $\ln D$. Then transform back to the original coordinates to obtain $T = U(\ln D)U^*$. As $\Lambda$ is real and also the eigenvalues of $\Lambda$ are real, the matrix $U$ of eigenvectors can also be chosen to be real. Thus the matrix $T$ is real. It is also hermitian, so it is symmetric, $T = T^{\mathrm{tr}}$.

Now define

$$\boxed{G(\xi) = \langle p, Tq \rangle} \ , \tag{5.21}$$

and with this choice, we find that

$$D_G = \langle Tp, \, \nabla_p \rangle - \langle Tq, \, \nabla_q \rangle \ . \tag{5.22}$$

Hence

$$-D_G\,q = Tq\;, \quad \text{and} \quad -D_G\,p = -Tp\;. \tag{5.23}$$

Then

$$Q = e^{-D_G}q = \sum_{j=0}^{\infty} \frac{T^j}{j!} q = e^T\,q = \Lambda\,q\;. \tag{5.24}$$

Likewise

$$P = e^{-D_G}p = \sum_{j=0}^{\infty} \frac{-T^j}{j!} p = e^{-T}\,q = \Lambda^{-1}p\;. \tag{5.25}$$

Therefore the general canonical scaling transformation can be written

$$\Xi = e^{-D_G}\xi = \begin{pmatrix} e^T & 0 \\ 0 & e^{-T} \end{pmatrix} \xi = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda^{-1} \end{pmatrix} \xi\;, \tag{5.26}$$

which has a Jacobian of the form (4.8).

**The case $\boldsymbol{\Lambda = R}$ is a real, orthogonal matrix:** We first find any matrix $T$ such that $R = e^T$. This can be done as follows: as $R$ is a real orthogonal, there is a unitary matrix $U$ so that $U^*RU = D$ is diagonal. A real orthogonal matrix is unitary, so its eigenvalues lie on the unit circle. Therefore $D = e^{i\Theta}$, for some real diagonal matrix $\Theta$, chosen so each of its eigenvalues $\theta_j$ corresponds to the eigenvalue $e^{i\theta_j}$ of $D$. Then

$$R = UDU^* = Ue^{i\Theta}\,U^* = e^{iU\Theta U^*}\;. \tag{5.27}$$

Thus we may take $T = iU\Theta U^*$.

As we assume that $R$ is a real orthogonal matrix,

$$R^{-1} = R^{\mathrm{tr}} = \left(e^T\right)^{\mathrm{tr}} = e^{T^{\mathrm{tr}}}\;. \tag{5.28}$$

We then define $D_G = \langle p, T\,\nabla_p\rangle - \langle q, T\,\nabla_q\rangle$ as in the previous case. But now

$$D_G = \left\langle T^{\mathrm{tr}}\,p,\, \nabla_p \right\rangle - \left\langle Tq,\, \nabla_q \right\rangle\;, \tag{5.29}$$

so

$$-D_G\,q = Tq\;, \quad \text{and} \quad -D_G\,p = -T^{\mathrm{tr}}\,p\;, \tag{5.30}$$

and

$$Q = e^{-D_G}q = \sum_{j=0}^{\infty} \frac{T^j}{j!} q = e^T\,q = R\,q\;. \tag{5.31}$$

Likewise

$$P = e^{-D_G}p = \sum_{j=0}^{\infty} \frac{(-T^{\mathrm{tr}})^j}{j!}\,p = e^{-T^{\mathrm{tr}}}\,p = \left(e^{-T}\right)^{\mathrm{tr}}\,p = R\,p\;. \tag{5.32}$$

Here we use $\left(e^{-T}\right)^{\mathrm{tr}} = \left(R^{-1}\right)^{\mathrm{tr}} = R$. Therefore the general canonical scaling transformation can be written

$$\Xi = e^{-D_G}\xi = \begin{pmatrix} e^T & 0 \\ 0 & \left(e^{-T}\right)^{\mathrm{tr}} \end{pmatrix} \xi = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix} \xi\;, \tag{5.33}$$

which has a Jacobian of the form (4.9).

**Theorem 5.1 (Hamilton Equations Yield Canonical Transformations).** *Let $G(\xi,\lambda)$ be a given function on phase space, and let $\xi(\lambda)$ be a solution* (2.39) *to Hamilton's equations* (2.33) *generated by $G$, with initial condition $\xi(0)=\xi$, and $\xi(\lambda)=\Xi$. Assume that $\xi(\lambda)$ is real analytic along the curve for all $\lambda$ in the interval $[\Lambda_1,\Lambda_2]\ni 0$. Then the family of transformation $\xi\mapsto\Xi$ is canonical for all $\lambda\in[\Lambda_1,\Lambda_2]$. This means*

$$[\Xi_i\,,\,\Xi_j]_\xi=\Gamma_{ij}\ ,\quad \textit{and}\quad J^{\mathrm{tr}}\Gamma J=\Gamma\ . \tag{5.34}$$

**Remark 5.2.** This result is presumably also true as long as a solution exists, without the assumption that $\xi(\lambda)$ is real analytic.

*Proof.* Let us compute the Jacobian $J_{D_G}$ of the transformation $D_G:\xi\mapsto D_G\xi$. In fact

$$D_G\,\xi_j=\sum_{k,\ell=1}^{2N}\frac{\partial G(\xi)}{\partial\xi_k}\,\Gamma_{k\ell}\,\frac{\partial}{\partial\xi_\ell}\,\xi_j=-\sum_{k=1}^{2N}\Gamma_{jk}\frac{\partial G(\xi)}{\partial\xi_k}\ . \tag{5.35}$$

Thus

$$(J_{D_G})_{ij}=\frac{\partial D_G\,\xi_j}{\partial\xi_i}=-\sum_{k=1}^{2N}\Gamma_{jk}\frac{\partial^2G(\xi,\lambda)}{\partial\xi_i\partial\xi_k}=-\left(\Gamma\mathfrak{h}\right)_{ji}=\left(\mathfrak{h}\Gamma\right)_{ij}\ , \tag{5.36}$$

where $\mathfrak{h}=\mathfrak{h}(\xi,\lambda)$ is the Hessian matrix of $G$ defined in (2.29).

Recall Proposition 3.7 in which we showed that composing transformations multiplies their Jacobian matrices. Furthermore, the Jacobian matrix of the sum of two transformations from the same initial point is the sum of the Jacobian matrices for the two transformations. So the Jacobian $J$ of $\xi_i(\lambda)=\left(e^{-\int_0^\lambda D_{G(\xi(\lambda_1),\lambda_1)}\,d\lambda_1}\right)_-\xi_i$ is given by

$$J=e^{-\lambda\mathfrak{h}\Gamma}\ ,\quad\text{with transpose}\quad J^{\mathrm{tr}}=e^{\lambda\Gamma\mathfrak{h}}\ . \tag{5.37}$$

Note that $\left(\Gamma\mathfrak{h}\right)^n\Gamma=\Gamma\left(\mathfrak{h}\Gamma\right)^n$. Hence $e^{\lambda\Gamma\mathfrak{h}}\,\Gamma=\Gamma e^{\mathfrak{h}\Gamma}$, and the Jacobian $J$ for the Taylor series of the solution satisfies

$$J^{\mathrm{tr}}\Gamma J=e^{\lambda\Gamma\mathfrak{h}}\,\Gamma e^{-\lambda\mathfrak{h}\Gamma}=\Gamma\,e^{\lambda\mathfrak{h}\Gamma}e^{-\lambda\mathfrak{h}\Gamma}=\Gamma\ . \tag{5.38}$$

Since we assume that the solution $\xi(\lambda)$ is analytic at $\lambda=0$, the Taylor series at $\lambda=0$ gives the solution away from $\lambda=0$. Hence this shows that $J$ is the actual Jacobian of the transformation $\xi\mapsto\Xi$, and (5.38) show that the transformation is canonical in a neighborhood of $\lambda=0$. We can repeat the same argument at any point on the solution curve. So as long as the solution is real analytic in $\lambda$, it determines a canonical transformation. □

## 6. Lie Derivatives and the Jacobi Identity

In this section we first study further the derivative $D_A$ that arises from considering the Poisson bracket $[A\,,\,B]_\xi$ as a transformation $D_A$ acting on $B$, namely $D_AB=[A\,,\,B]_\xi$. In the study of differential geometry such a derivative $D_A$ is called a "Lie derivative" after the Norwegian mathematician Sophus Lie (1842–1899). He was the person who initiated the study of continuous symmetry groups, and today they are known as "Lie groups." Lie became the successor to Felix Klein (1849–1925) in 1886, when Klein left Leipzig to become professor in Göttingen. Lie suffered from anemia, due to a deficiency of vitamin absorption. He ultimately had to resign, and he died young at age 56.

Lie derivatives have a number of interesting properties that are relevant for classical mechanics. The first four of these (6.1)–(6.4) are elementary, and they follow immediately from the definition of the Poisson bracket. Nevertheless, it is interesting to state them in terms of the derivative. The fifth property that we give requires some work to check that it holds, so we state it as Proposition 6.1. After establishing this identity, we translate it back into an identity, known as the Jacobi identity for Poisson brackets, in Proposition 6.2. We motivate this by explaining the Jacobi identity for matrix multiplication. Other Jacobi identities arise in gauge theory (in particular in Yang-Mills theory), in differential geometry, and hence also in general relativity. In mathematics the Jacobi identity is fundamental in developing the theory of Lie algebras.

### 6.1. **Basic Identities for Derivatives.**

- The derivative $D_A$ is linear in $A$. For constants $\lambda_1, \lambda_2$ and functions $A_1, A_2$,

$$\lambda_1 D_{A_1} + \lambda_2 D_{A_2} = D_{\lambda_1 A_1 + \lambda_2 A_2} \ . \tag{6.1}$$

- The derivative $D_A$ acts linearly on $B$. For constants $\lambda_1, \lambda_2$ and functions $B_1, B_2$,

$$D_A(B_1 + B_2) = D_A\, B_1 + D_A\, B_2 \ . \tag{6.2}$$

- The derivative $D_A$ satisfies the product rule,

$$D_A(BC) = (D_A B)C + B(D_A C) \ . \tag{6.3}$$

- The derivative satisfies

$$D_A B = -D_B A \ . \quad \text{In particular} \quad D_A A = 0 \ . \tag{6.4}$$

  This is a consequence of $[A\,,\,B]_\xi = -\,[B\,,\,A]_\xi$.
- The most important property for the derivative $D_A$ is a formula for the commutator of the derivative $D_A$ with another derivative $D_B$, namely a beautiful expression for $[D_A\,,\,D_B] = D_A D_B - D_B D_A$. We state this fundamental identity as a proposition.

**Proposition 6.1 (The Fundamental Commutator Formula).** *For twice differentiable $A, B$, the identity*

$$\boxed{[D_A\,,\,D_B] = D_{[A\,,\,B]_\xi}} \ . \tag{6.5}$$

*holds. Also* (6.5) *is equivalent to the statement that for any twice-differentiable $C(\xi)$,*

$$\boxed{D_A\, D_B\, C - D_B\, D_A\, C = D_{[A\,,\,B]_\xi}\, C} \ . \tag{6.6}$$

*Proof.* Let us first establish (6.6). We compute the left side of (6.6). We claim that the terms involving second derivatives of $C$ cancel in the difference, leaving only the terms that are first derivatives of $C$. To see this, note that $\frac{\partial^2 C}{\partial\xi_j\partial\xi_l} = \frac{\partial^2 C}{\partial\xi_l\partial\xi_j}$, so one can write the terms involving second derivatives of $C$ in the expression $D_A\, D_B\, C$ as

$$\sum_{i,j,k,l=1}^{2N} \Gamma_{ij}\,\Gamma_{kl}\,\frac{\partial A}{\partial\xi_i}\,\frac{\partial B}{\partial\xi_k}\,\frac{\partial^2 C}{\partial\xi_j\partial\xi_l} = \sum_{i,j,k,l=1}^{2N} \Gamma_{il}\,\Gamma_{kj}\,\frac{\partial A}{\partial\xi_i}\,\frac{\partial B}{\partial\xi_k}\,\frac{\partial^2 C}{\partial\xi_j\partial\xi_l} = \sum_{i,j,k,l=1}^{2N} \Gamma_{kl}\,\Gamma_{ij}\,\frac{\partial A}{\partial\xi_k}\,\frac{\partial B}{\partial\xi_i}\,\frac{\partial^2 C}{\partial\xi_j\partial\xi_l}\,.$$

The last expression on the right is also equal to the terms with second derivatives of $C$ in the expression $D_B\, D_A\, C$. Therefore the second derivatives of $C$ cancel in the commutator $[D_A\,,\, D_B]\, C$.

The remaining terms in (6.6) all involve first derivatives of $C$. Use the product rule for differentiation to write them as

$$\begin{aligned}
D_A\, D_B\, C - D_B\, D_A\, C &= \sum_{i,j=1}^{2N}\sum_{k,l=1}^{2N}\left(\frac{\partial A}{\partial\xi_i}\Gamma_{ij}\frac{\partial^2 B}{\partial\xi_j\partial\xi_k} - \frac{\partial B}{\partial\xi_i}\Gamma_{ij}\frac{\partial^2 A}{\partial\xi_j\partial\xi_k}\right)\Gamma_{kl}\frac{\partial C}{\partial\xi_l}\\
&= \sum_{i,j=1}^{2N}\sum_{k,l=1}^{2N}\left(\frac{\partial A}{\partial\xi_i}\Gamma_{ij}\frac{\partial^2 B}{\partial\xi_j\partial\xi_k} + \frac{\partial^2 A}{\partial\xi_j\partial\xi_k}\Gamma_{ji}\frac{\partial B}{\partial\xi_i}\right)\Gamma_{kl}\frac{\partial C}{\partial\xi_l}\\
&= \sum_{i,j=1}^{2N}\sum_{k,l=1}^{2N}\left(\frac{\partial A}{\partial\xi_i}\Gamma_{ij}\frac{\partial^2 B}{\partial\xi_j\partial\xi_k} + \frac{\partial^2 A}{\partial\xi_i\partial\xi_k}\Gamma_{ij}\frac{\partial B}{\partial\xi_j}\right)\Gamma_{kl}\frac{\partial C}{\partial\xi_l}\\
&= \sum_{k,l=1}^{2N}\left(\frac{\partial}{\partial\xi_k}\,[A\,,\,B]_\xi\right)\Gamma_{kl}\frac{\partial C}{\partial\xi_l}\\
&= D_{[A\,,\,B]_\xi}\, C\,.
\end{aligned}$$

This is the case for all $C$, so the desired identity (6.6) holds.

In order to complete the proof of the proposition, we need to show the equivalence of (6.6) with (6.5). Let $K(\xi) = [D_A\,,\, D_B] - [A\,,\,B]_\xi$. Clearly if $K = 0$, which is (6.5), then $D_K C = 0$, which is (6.6). So we need only establish the converse.

Assume $D_K C = 0$ for arbitrary $C$. If $K$ is a constant, then $D_K = 0$ as desired. We now rule out the case that $K$ is not a constant. We show that if that were true, we could choose $C$ such that $D_K C \neq 0$.

In case $K$ is not constant, for at least one of the $\xi$ coordinates, say $\xi_\ell$, in the neighborhood of some point in phase space, both $\frac{\partial K}{\partial\xi_\ell} \neq 0$ and also $K \neq 0$. This means that $K\frac{\partial K}{\partial\xi_\ell} \neq 0$. Now choose $C = K\sum_{j=1}^{2N}\Gamma_{\ell j}\xi_j$, and compute $D_K C$. In order to do this, use the properties

(6.2)–(6.4), and also the identity $\Gamma\Gamma^{\rm tr} = I$. The answer is $K\frac{\partial K}{\partial \xi_\ell}$, which one sees as follows:

$$\begin{aligned} D_K C &= D_K\left(K\sum_{j=1}^{2N}\Gamma_{\ell j}\xi_j\right) = K\sum_{j=1}^{2N}\Gamma_{\ell j}D_K\xi_j = K\sum_{j,r,s=1}^{2N}\Gamma_{\ell j}\frac{\partial K}{\partial \xi_r}\Gamma_{rs}\frac{\partial \xi_j}{\partial \xi_s} \\ &= K\sum_{j,rs,=1}^{2N}\Gamma_{\ell j}\frac{\partial K}{\partial \xi_r}\Gamma_{rs}\delta_{j,s} = K\sum_{j,r=1}^{2N}\Gamma_{\ell j}\frac{\partial K}{\partial \xi_r}\Gamma^{\rm tr}_{jr} = K\sum_{r=1}^{2N}(\Gamma\Gamma^{\rm tr})_{\ell j}\frac{\partial K}{\partial \xi_r} \\ &= K\frac{\partial K}{\partial \xi_\ell}\,. \end{aligned} \tag{6.7}$$

Since $K\frac{\partial K}{\partial \xi_\ell} \neq 0$, this choice of $K$ and $C$ is incompatible with $D_K C = 0$. Thus we have ruled out the case that $K$ is not a constant, and we conclude that $D_K = 0$. □

6.2. **The Jacobi Identity for Matrices.** The name "Jacobi identity" arises from a corresponding identity for matrix multiplication. One can check that for three $N \times N$ matrices $A, B, C$, it is always the case that the commutator $[A\,,\,B] = AB - BA$ satisfies

$$[[A\,,\,B]\,,\,C] + [[B\,,\,C]\,,\,A] + [[C\,,\,A]\,,\,B] = 0\,. \tag{6.8}$$

The three terms are cyclic permutations of the double commutators, and they add up to zero. This is the *Jacobi identity for matrix multiplication.*

One can check the Jacobi identity by writing out the four terms in each of the three double commutators. For example, the first term is $[[A\,,\,B]\,,\,C] = [A\,,\,B]\,C - C\,[A\,,\,B] = ABC - BAC - CAB + CBA$. Thus the Jacobi identity involves a total of twelve terms. Each of the $3! = 6$ possible orders for $A, B, C$ occurs exactly twice among the twelve terms: once with a positive sign and once with a negative sign. Explicitly

$$\begin{aligned} &[[A\,,\,B]\,,\,C] + [[B\,,\,C]\,,\,A] + [[C\,,\,A]\,,\,B] \\ &\quad = ABC - BAC - CAB + CBA + BCA - CBA \\ &\qquad\quad -ABC + ACB + CAB - ACB - BCA + BAC \\ &\quad = 0\,. \end{aligned}$$

6.3. **The Jacobi Identity for Poisson Brackets.** We have seen that the derivative $D_A$ acting on a function $C$ is the Poisson bracket $D_A C = [A\,,\,C]_\xi$ on phase space. This bracket has many of the same algebraic properties as the commutator of matrices $[A\,,\,C]$. Thus one might guess that the Poisson bracket could satisfy a Jacobi identity similar to (6.8). We now show that is the case, with an argument that reduces to the identity already established in Proposition 6.1.

**Proposition 6.2.** *For twice differentiable functions $A, B, C$ on phase space, the Jacobi identity for Poisson brackets holds, namely*

$$\Big[[A\,,\,B]_\xi\,,\,C\Big]_\xi + \Big[[B\,,\,C]_\xi\,,\,A\Big]_\xi + \Big[[C\,,\,A]_\xi\,,\,B\Big]_\xi = 0\,. \tag{6.9}$$

*An equivalent Jacobi identity follows from the anti-symmetry of the Poisson bracket,*

$$\Big[A\,,\,[B\,,\,C]_\xi\Big]_\xi+\Big[B\,,\,[C\,,\,A]_\xi\Big]_\xi+\Big[C\,,\,[A\,,\,B]_\xi\Big]_\xi=0\;. \tag{6.10}$$

*Proof.* Use the relation $[X\,,\,C]_\xi = D_X C$ to rewrite the first term in (6.9) as $D_{[A\,,\,B]_\xi}C$. However the commutator identity (6.5) of Proposition 6.1 applied to the $D_{[A\,,\,B]_\xi}$ shows this to be

$$\Big[[A\,,\,B]_\xi\;,\;C\Big]_\xi=[D_A\,,\,D_B]\,C\;. \tag{6.11}$$

The Poisson bracket is skew symmetric, so the second term in (6.9) can be written

$$\Big[[B\,,\,C]_\xi\;,\;A\Big]_\xi=-\Big[A\,,\,[B\,,\,C]_\xi\Big]_\xi=-D_AD_BC\;. \tag{6.12}$$

Likewise the third term in (6.9) can be written,

$$\Big[[C\,,\,A]_\xi\;,\;B\Big]_\xi=-\Big[B\,,\,[C\,,\,A]_\xi\Big]_\xi=\Big[B\,,\,[A\,,\,C]_\xi\Big]_\xi=D_BD_AC\;. \tag{6.13}$$

Adding (6.11)–(6.13) we have

$$\Big[[A\,,\,B]_\xi\;,\;C\Big]_\xi+\Big[[B\,,\,C]_\xi\;,\;A\Big]_\xi+\Big[[C\,,\,A]_\xi\;,\;B\Big]_\xi=[D_A\,,\,D_B]\,C-[D_A\,,\,D_B]\,C=0\;,$$

to complete the proof. □

## 7. Canonical Transformations and Symmetry

The Hamilton equations give the most beautiful relation between conservation laws and symmetry. Basically we study two different groups. One group is the dynamical group characterized by a Hamiltonian $H(\xi)$ and the corresponding flow on phase space given by the solution to Hamilton's equations, $\dot{\xi}(t)=-\,[H\,,\,\xi]_{\xi(t)}$. As we have seen in Theorem 5.1, this gives rise to a family of canonical transformations,

$$\xi\mapsto\Xi=\xi(t)\;. \tag{7.1}$$

We can also write this family of transformations as a one-parameter time translation group $T_t$, in which case $T_tT_s=T(t+s)$, and $T(0)=Id$. We can write

$$\xi(t+s)=T(t)\xi(s)\;. \tag{7.2}$$

(Note that in case $H$ is explicitly time-dependent, there is a more-general way to obtain a family of canonical transformations that solve Hamilton's equations, but in this case the formula for the solution involves Dyson's time-ordered product.)

We also consider a second group of canonical transformations that arise from a different set of Hamilton's equations, generated by one or more functions $G_1(\xi),\ldots,G_r(\xi)$ on phase space, each of which have zero Poisson bracket with the Hamiltonian, namely

$$[H\,,\,G_i]_\xi=0\;,\quad\text{for}\quad i=1,\ldots,r\;. \tag{7.3}$$

The second group $\mathcal{G}$ is generated by products of the various one-parameter groups

$$V_i(\lambda_i) = e^{-\lambda_i D_{G_i}} \; . \tag{7.4}$$

The simplest case is to have one generator $G = G_1$, and $V_\lambda = e^{-\lambda D_G}$ is a one-parameter group. For example, if $N = 2$ and $G = q_1p_2 - q_2p_1$, then

$$-\,D_G = -\sum_{i,j=1}^{4} \frac{\partial G}{\partial q_i}\Gamma_{ij}\frac{\partial}{\partial p_j} = \left(p_1\frac{\partial}{\partial p_2} - p_2\frac{\partial}{\partial p_1}\right) + \left(q_1\frac{\partial}{\partial q_2} - q_2\frac{\partial}{\partial q_1}\right) \; . \tag{7.5}$$

Then $V_\lambda$ is the one-parameter group of simultaneous rotations by angle $\lambda$ in the $(x_1, x_2)$ plane and in the $(p_1, p_2)$ plane. The explicit form of this transformation on $\mathbb{R}^4$ is given by the linear transformation,

$$\xi(\lambda) = e^{-\lambda D_G}\xi = \begin{pmatrix} \cos\lambda & -\sin\lambda & 0 & 0 \\ \sin\lambda & \cos\lambda & 0 & 0 \\ 0 & 0 & \cos\lambda & -\sin\lambda \\ 0 & 0 & \sin\lambda & \cos\lambda \end{pmatrix} \xi \; , \quad \text{where } \xi = \begin{pmatrix} q_1 \\ q_2 \\ p_1 \\ p_2 \end{pmatrix} \; . \tag{7.6}$$

The action of $D_G$ is the "infinitesimal" form of this transformation

$$-\,D_G\,\xi(\lambda) = \frac{d}{d\lambda}\,\xi(\lambda) = \begin{pmatrix} -\sin\lambda & -\cos\lambda & 0 & 0 \\ \cos\lambda & -\sin\lambda & 0 & 0 \\ 0 & 0 & -\sin\lambda & -\cos\lambda \\ 0 & 0 & \cos\lambda & -\sin\lambda \end{pmatrix} \xi \; , \tag{7.7}$$

so

$$-\,D_G\,\xi = \frac{d}{d\lambda}\,\xi(\lambda)\bigg|_{\lambda=0} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \xi = \begin{pmatrix} -q_2 \\ q_1 \\ -p_2 \\ p_1 \end{pmatrix} \; . \tag{7.8}$$

We will give many other examples shortly.

The connection between a conserved quantity $G$ and a symmetry is now encapsulated by the results from earlier in these notes. We summarize the bottom line in the following:

**Theorem 7.1** **(Conservation and Symmetry).** *Let $G$ be conserved by the Hamiltonian $H$. In other words $[H\,,\,G]_\xi = D_HG = 0$. Then the group $V_\lambda = e^{-\lambda D_G}$ leaves $H$ invariant, namely $e^{-\lambda D_G}H = H$. Also $T_t = e^{-tD_H}$ satisfies $T_tV_\lambda = V_\lambda T_t$ for all $\lambda, t$. Thus the group $V_\lambda$ transforms one solution $T_t\xi$ of Hamilton's time evolution equation into a family of other solutions $T_tV_\lambda\xi = V_\lambda T_t\xi$.*

*Proof.* If we evaluate $G(\xi(t))$ along an orbit $\xi(t) = T_t\xi = e^{-t\mathcal{D}_H}\xi$, then we claim that

$$\frac{dG(\xi(t))}{dt} = \sum_i \frac{\partial G\xi(t))}{\partial \xi_i(t)}\frac{d\xi_i(t)}{dt} = \sum_{ij,i} \frac{\partial G\xi(t))}{\partial \xi_i(t)}\Gamma_{ij}\frac{\partial H(\xi(t))}{\partial \xi_j(t)} = -D_HG(\xi(t)) \; .$$

Thus $G(\xi(t)) = e^{-tD_H} G(\xi) = T_t G(\xi)$ and

$$G(\xi(t)) = T_t\, G(\xi) = \sum_{n=0}^{\infty} \frac{(-\lambda D_H)^n}{n!} G(\xi) = G(\xi)\ . \tag{7.9}$$

As $[D_G\, ,\, D_H] = D_{[G\, ,\, H]_\xi} = 0$, it follows that $V_\lambda T_t = T_t V_{\leq}$. In particular, if the groups $T_t$ and $V_\lambda$ are real analytic in $t$ and $\lambda$, we can rearrange the power series in $t$ and in $\lambda$ to show,

$$T_t V_\lambda = V_\lambda T_t\ . \tag{7.10}$$

Apply this identity to the initial point $\xi$ in phase space. Thus one sees that $V_\lambda$ transforms the solution to Hamilton's equations in time, $T_t \xi$, into the solution $T_t V_\lambda \xi$, where the symmetry $V_\lambda$ acts on the initial condition. □

**Remark 7.2.** It is important to remark that commutativity of the two groups of canonical transformations

$$V_\lambda T_t = T_t V_\lambda\ , \tag{7.11}$$

does not mean that their generators $D_G$ and $D_H$ commute. We only can conclude that $[D_G\, ,\, D_H] = D_{[G\, ,\, H]_\xi} = 0$. This will be the case as long as $[G\, ,\, H]_\xi$ is a constant, and the constant may not be zero. Consider the case N=1, $G = q$, and $H = p$. Then $[q\, ,\, p]_\xi = 1$, but $[D_q\, ,\, D_p] = D_{[q\, ,\, p]_\xi} = 0$.

## 8. More Details of Some Canonical Transformations

Let us begin with Table 1 that summarizes a few symmetries arising from canonical transformations generated by a function $G(\xi)$ on phase space, or a collection of such functions, giving rise to one or more Hamilton's equations; here we call the $G$'s "Hamiltonians." We analyze a few of these situations in detail.

8.1. **Angular Momentum.** Consider a single particle in 3-dimensional configuration space $\vec{x} \in \mathbb{R}^3$, with the corresponding momentum space $\vec{p} \in \mathbb{R}^3$. The angular momentum about the origin is

$$\vec{L} = \vec{x} \times \vec{p} = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \begin{pmatrix} x_2 p_3 - x_3 p_2 \\ x_3 p_1 - x_1 p_3 \\ x_1 p_2 - x_2 p_1 \end{pmatrix}\ . \tag{8.1}$$

Here the cross product can also be written in terms of components as

$$(\vec{x} \times \vec{p})_i = \sum_{j,k=1}^{3} \epsilon_{ijk}\, x_j\, p_k\ . \tag{8.2}$$

We introduce here the totally anti-symmetric tensor $\epsilon_{ijk}$ depending on three indices $i, j, k = 1, 2, 3$, and having the property $\epsilon_{123} = 1$. In other words,

$$\epsilon_{ijk} = \begin{cases} 1\ , & \text{if (i,j,k) is an even permutation of } (1,2,3) \\ -1\ , & \text{if (i,j,k) is an odd permutation of } (1,2,3) \\ 0\ , & \text{otherwise} \end{cases}\ . \tag{8.3}$$

**Examples of the Action of $D_G$ on the Coordinates $q$ and $p$.**

| $G(\xi)$ | $D_G$ | $Q_j = e^{-\lambda D_G}\, q_j$ | $P_j = e^{-\lambda D_G}\, p_j$ |
|---|---|---|---|
| $q_i$ | $\frac{\partial}{\partial p_i}$ | $q_j$ | $p_j - \delta_{ij}\,\lambda$ |
| $p_i$ | $-\frac{\partial}{\partial q_i}$ | $q_j + \delta_{ij}\,\lambda$ | $p_j$ |
| $\sum_{i=1}^N \Omega_i q_i\, p_i$ | $\sum_{i=1}^N \Omega_i \left(p_i\,\frac{\partial}{\partial p_i} - q_i\,\frac{\partial}{\partial q_i}\right)$ | $e^{\lambda\Omega_j}\, q_j$ | $e^{-\lambda\Omega_j}\, p_j$ |
| $\frac{1}{2}\sum_{i=1}^N \left(p_i^2 + \omega_i^2 q_i^2\right)$ | $\sum_{i=1}^N \left(\omega_i^2 q_i \frac{\partial}{\partial p_i} - p_i \frac{\partial}{\partial q_i}\right)$ | $q_j \cos(\omega_j\lambda) + p_j\omega_j^{-1}\sin(\omega_j\lambda)$ | $p_j\cos(\omega_j\lambda) - q_j\omega_j\sin(\omega_j\lambda)$ |
| $L_3 = q_1 p_2 - q_2 p_1$ | $D_{L_3} = p_2\,\frac{\partial}{\partial p_1} - p_1\,\frac{\partial}{\partial p_2}$ $-q_1\,\frac{\partial}{\partial q_2} + q_2\,\frac{\partial}{\partial q_1}$ | $Q_1 = q_1\cos\lambda - q_2\sin\lambda$<br>$Q_2 = q_1\sin\lambda + q_2\cos\lambda$<br>$Q_3 = q_3$ | $P_1 = p_1\cos\lambda - p_2\sin\lambda$<br>$P_2 = p_1\sin\lambda + p_2\cos\lambda$<br>$P_3 = p_3$ |
| $\vec{p}_{\text{tot}} = \sum_{i=1}^n \vec{a}\cdot\vec{p}^{\,(i)}$ | $-\sum_{i=1}^n \vec{a}\cdot\nabla_{\vec{q}^{\,(i)}}$ | $\vec{q}^{\,(j)} + \lambda\vec{a}$ | $\vec{p}^{\,(j)}$ |
| $\epsilon = \left(1 + \frac{2HL^2}{mk^2}\right)^{1/2}$ | (Kepler $H$) | Complicated | Complicated |

TABLE 1. Canonical Transformations from Hamiltonians of the form $G(\xi)$.

Then one can find the Poisson bracket between $L_1$ and $L_2$,

$$\begin{aligned} [L_1\,,\, L_2]_\xi &= [x_2p_3 - x_3p_2\,,\, x_3p_1 - x_1p_3]_\xi \\ &= [x_2p_3\,,\, x_3p_1]_\xi - [x_2p_3\,,\, x_1p_3]_\xi - [x_3p_2\,,\, x_3p_1]_\xi + [x_3p_2\,,\, x_1p_3]_\xi \\ &= -x_2p_1 - 0 - 0 + p_2x_1 \\ &= L_3\,. \end{aligned} \tag{8.4}$$

Similarly,

$$[L_2\,,\, L_3]_\xi = L_1\,, \quad \text{and} \quad [L_3\,,\, L_1]_\xi = L_2\,. \tag{8.5}$$

Clearly these three Poisson brackets determine the nine Poisson brackets $[L_i\,,\, L_j]_\xi$ for $i,j = 1,2,3$, using the skew symmetry of the Poisson bracket. The nine relations can also be written as one equation,

$$\boxed{[L_i\,,\, L_j]_\xi = \sum_{k=1}^{3} \epsilon_{ijk} L_k}\,. \tag{8.6}$$

8.1.1. *$n$ Particles.* In the case of a set of $n$ particles in three-space, let us index the positions by $\vec{x}^{\,(\alpha)}$, for $\alpha = 1,\ldots,n$. Denote the corresponding momenta by $\vec{p}^{\,(\alpha)}$, for $\alpha = 1,\ldots,n$.

Then the total angular momentum about the origin is

$$\vec{L}^{\,\mathrm{tot}} = \sum_{\alpha=1}^{n} \vec{L}^{\,(\alpha)} = \sum_{\alpha=1}^{n} \vec{x}^{\,(\alpha)} \times \vec{p}^{\,(\alpha)} \,. \tag{8.7}$$

Then we can compute the Poisson bracket relations between the components of the total angular momentum, and we find that the algebraic structure is the same as in the case of the angular momentum for a single particle,

$$\boxed{\left[L_i^{\mathrm{tot}}\,,\, L_j^{\mathrm{tot}}\right]_\xi = \sum_{k=1}^{3} \epsilon_{ijk}\, L_k^{\mathrm{tot}}}\,, \tag{8.8}$$

as

$$\left[L_i^{\mathrm{tot}}\,,\, L_j^{\mathrm{tot}}\right]_\xi = \sum_{\alpha,\beta=1}^{n} \left[L_i^{(\alpha)}\,,\, L_j^{(\beta)}\right]_\xi = \sum_{\alpha=1}^{n} \left[L_i^{(\alpha)}\,,\, L_j^{(\alpha)}\right]_\xi = \sum_{\alpha=1}^{n}\sum_{k=1}^{3} \epsilon_{ijk}\, L_k^{(\alpha)} = \sum_{k=1}^{3} \epsilon_{ijk}\, L_k^{\mathrm{tot}} \,. \tag{8.9}$$

8.2. **Angular Momentum and the group $SO(3)$ of Rotations on $\mathbb{R}^3$.** A rotation of a vector in 3-space is determined by a $3\times 3$ real, orthogonal matrix $R$, namely a matrix for which $R^{\mathrm{tr}}R = I$. The rotation is *proper* in case $\det R = +1$, rather than $\det R = -1$. The group of proper rotations is called $SO(3)$.

In the case of a single particle in 3-space, the angular momentum vector $\vec{L}$ is intimately connected with the group of proper rotations of the coordinates, both the configuration space coordinate $\vec{q}$ and the momentum coordinate $\vec{p}$. For a set of $n$ particles, the total angular momentum vector $\vec{L}^{\,\mathrm{tot}}$ plays the same role for the simultaneous rotation of each particle in the set; it is related to the same rotation for each $\vec{q}^{\,(j)}$ and each $\vec{p}^{\,j}$, for $j = 1, \ldots, n$.

Given a unit vector $\vec{n}$ in 3-space, the component of $\vec{L}$ in the direction $\vec{n}$, is the generator of rotations about $\vec{n}$. For a single particle, this is just the function $G = \vec{L}\cdot\vec{n}$, and we now show that

$$e^{-\theta\, D_G} = e^{-\theta D_{\vec{L}\cdot\vec{n}}} \tag{8.10}$$

is the canonical transformation giving a rotation by angle $\theta$ about the axis $\vec{n}$.

8.2.1. *Rotation of Phase Space Coordinates.* We see the action of $D_{\vec{L}\cdot\vec{n}}$ by computing what it looks like. In fact

$$D_{\vec{L}\cdot\vec{n}} = \sum_{j=1}^{3} n_j\, D_{L_j}\ ,$$

where

$$D_{L_1} = -\left(x_2\,\frac{\partial}{\partial x_3} - x_3\,\frac{\partial}{\partial x_2}\right) - \left(p_2\,\frac{\partial}{\partial p_3} - p_3\,\frac{\partial}{\partial p_2}\right)\ ,$$

$$D_{L_2} = -\left(x_3\,\frac{\partial}{\partial x_1} - x_1\,\frac{\partial}{\partial x_3}\right) - \left(p_3\,\frac{\partial}{\partial p_1} - p_1\,\frac{\partial}{\partial p_3}\right)\ ,$$

and

$$D_{L_3} = -\left(x_1\,\frac{\partial}{\partial x_2} - x_2\,\frac{\partial}{\partial x_1}\right) - \left(p_1\,\frac{\partial}{\partial p_2} - p_2\,\frac{\partial}{\partial p_1}\right)\ .$$

Then one sees that

$$[L_i\,,\,x_j]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}\,x_k = D_{L_i}\,x_j\ ,\quad\text{and}\quad [L_i\,,\,p_j]_\xi = D_{L_i}\,p_j = \sum_{k=1}^{3}\epsilon_{ijk}\,p_k\ . \tag{8.11}$$

Thus

$$\left[\vec{L}\cdot\vec{n}\,,\,x_j\right] = D_{\vec{L}\cdot\vec{n}}\,x_j = \sum_{i,k=1}^{3}\epsilon_{ijk}\,n_i\,x_k = -\left(\vec{n}\times\vec{x}\right)_j\ , \tag{8.12}$$

and

$$\left[\vec{L}\cdot\vec{n}\,,\,p_j\right] = D_{\vec{L}\cdot\vec{n}}\,p_j = \sum_{i,k=1}^{3}\epsilon_{ijk}\,n_i\,p_k = -\left(\vec{n}\times\vec{p}\right)_j\ . \tag{8.13}$$

One can express these identities as the vector relations,

$$D_{\vec{L}\cdot\vec{n}}\,\vec{x} = -\vec{n}\times\vec{x}\ ,\quad\text{and}\quad D_{\vec{L}\cdot\vec{n}}\,\vec{p} = -\vec{n}\times\vec{p}\ . \tag{8.14}$$

Thus we can sum the exponential series to find that on the configuration-space coordinate $\vec{x}$, or the momentum coordinate $\vec{p}$ one has

$$\vec{X} = e^{-\theta D_{\vec{L}\cdot\vec{n}}}\ \vec{x} = R(\theta,\vec{n})\,\vec{x}\ ,\quad\text{and}\quad \vec{P} = e^{-\theta D_{\vec{L}\cdot\vec{n}}}\ \vec{p} = R(\theta,\vec{n})\,\vec{p}\ . \tag{8.15}$$

The group of transformations generated by the different $\vec{L}\cdot\vec{n}$ for different vectors $\vec{n}$ is called the group $SO(3)$, namely the special orthogonal group of $3\times 3$ matrices. These are the $3\times 3$ orthogonal matrices with the "special" value of their determinant, namely $+1$.

On the phase for several particles, the corresponding rotation transformation is $e^{-\theta D_{\vec{L}^{\,\rm tot}\cdot\vec{n}}}$ in place of $e^{-\theta D_{\vec{L}\cdot\vec{n}}}$. We go into these details about the symmetries of 3-space and of 4-space-time in another set of notes on rotations and Lorentz transformations.

8.3. **General "Vectors" under Rotations in 3-Space.** One can abstract the notion of a 3-*vector* on phase space as a set of three functions $\vec{v}$ with components $\vec{v} = (v_1, v_2, v_3)$ that satisfy some important relations.

**Definition 8.1.** *A set of three functions $\vec{v} = (v_1, v_2, v_3)$ on phase space is a 3-vector under rotations, if it satisfies the Poisson bracket relations:*

$$\boxed{[L_i\,,\,v_j]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}\,v_k\ ,\quad \textit{for all } i,j = 1,2,3}\ . \tag{8.16}$$

*A function $s(\xi)$ on phase space is a scalar under rotations, if*

$$\boxed{[L_i\,,\,s]_\xi = 0\ ,\quad \textit{for } i = 1,2,3}\ . \tag{8.17}$$

Note that by the skew symmetry of $\epsilon_{ijk}$ this is the same as requiring that

$$[v_i\,,\,L_j]_\xi = -\sum_{k=1}^{3} \epsilon_{ijk}\, v_k\;, \quad \text{for all} \quad i,j = 1,2,3\;. \tag{8.18}$$

Thus for any 3-vector under rotations,

$$\boxed{[L_i\,,\,v_j]_\xi = -\,[L_j\,,\,v_i]_\xi}\;. \tag{8.19}$$

Let $\vec{n}$ be any constant unit vector. Then (8.16) holds, if and only if

$$\boxed{\left[\vec{L}\cdot\vec{n}\,,\,v_j\right]_\xi = -(\vec{n}\times\vec{v})_j \quad \text{for all constant unit vectors } \vec{n}}\;. \tag{8.20}$$

So we can also take (8.20) in place of Definition 8.1 as determining a 3-vector under rotations. For any scalar function $A$ on phase space,

$$\boxed{e^{-\theta D_{\vec{L}\cdot\vec{n}}} A = A}\;. \tag{8.21}$$

For any vector function $\vec{v}$, on phase space,

$$\boxed{e^{-\theta D_{\vec{L}\cdot\vec{n}}}\,\vec{v} = R(\theta,\vec{n})\vec{v}}\;, \tag{8.22}$$

where $R(\theta,\vec{n}) \in SO(3)$ is the real, orthogonal matrix giving a rotation by angle $\theta$ about the axis $\vec{n}$.

In the following we use the vector product $\vec{v}\times\vec{w}$ of two 3-vectors $\vec{v}$ and $\vec{w}$, as the object with components,

$$\left(\vec{v}\times\vec{w}\right)_i = \sum_{j,k=1}^{3} \epsilon_{ijk} v_j w_k\;. \tag{8.23}$$

Here $\epsilon_{ijk}$ is the totally antisymmetric symbol,

$$\epsilon_{ijk} = \begin{cases} 1, & \text{if } i,j.k \text{ is an even permutation of 1,2,3} \\ -1, & \text{if } i,j.k \text{ is an odd permutation of 1,2,3} \\ 0, & \text{otherwise} \end{cases} \tag{8.24}$$

**Proposition 8.2** (**Some Properties of Scalars and 3-Vectors under Rotations).** *Let $s$ be a scalar function on phase space under rotations, and let $\vec{v}$ and $\vec{w}$ be 3-vectors under rotations. Then:*

*i. The product $s\vec{v}$ is a 3-vector under rotations*
*ii. The sum $\vec{v}+\vec{w}$ is a 3-vector under rotations.*
*iii. The scalar product $\vec{v}\cdot\vec{w} = \langle\vec{v}\cdot\vec{w}\rangle = \sum_{i=1}^{3} v_i w_i$ is a scalar under rotations.*
*iv. The vector product $(\vec{v}\times\vec{w})_i = \sum_{j,k=1}^{3} \epsilon_{ijk}\, v_j w_k$ is a 3-vector under rotations.*

*Proof.* **Part i:** By the product rule for Poisson brackets (2.5), one has

$$[L_i\,,\,sv_j]_\xi = \sum_{j=1}^{3} [L_i\,,\,s]_\xi\, v_j + [L_i\,,\,v_j]_\xi\, s = \sum_{j=1}^{3} [L_i\,,\,v_j]_\xi\, s = \sum_{j=1}^{3} \epsilon_{ijk} s v_k\;. \tag{8.25}$$

**Part ii:** This is clear from linearity of the relation for a 3-vector.
**Part iii:** By the product rule for Poisson brackets (2.5), one has

$$[L_i\,,\,\vec{v}\cdot\vec{w}]_\xi = \sum_{j=1}^{3}[L_i\,,\,v_j]_\xi\,w_j + [L_i\,,\,w_j]_\xi\,v_j = \sum_{j,k=1}^{3}\epsilon_{ijk}(v_k w_j + w_k v_j) = 0\,. \tag{8.26}$$

Here we use the fact that the expression in parentheses is symmetric under the interchange of $j$ and $k$, while $\epsilon_{ijk}$ is skew symmetric. Hence the sum vanishes and $\vec{v}\cdot\vec{w}$ is a scalar.
**Part iv:** For $\vec{v}\times\vec{w}$, we use the product rule for Poisson brackets to show for any constant unit 3-vector $\vec{n}$,

$$\begin{aligned}\left[\vec{L}\cdot\vec{n}\,,\,\vec{v}\times\vec{w}\right]_\xi &= \left[\vec{L}\cdot\vec{n}\,,\,\vec{v}\right]_\xi\times\vec{w} + \vec{v}\times\left[\vec{L}\cdot\vec{n}\,,\,\vec{w}\right]_\xi \\ &= -\left(\vec{n}\times\vec{v}\right)\times w - \vec{v}\times\left(\vec{n}\times\vec{w}\right) = -\vec{n}\times\left(\vec{v}\times\vec{w}\right)\,. \end{aligned}\tag{8.27}$$

In the last equality we use the rule for products of 3-vectors,

$$\vec{n}\times(\vec{v}\times\vec{w}) = (\vec{n}\cdot\vec{w})\vec{v} - (\vec{n}\cdot\vec{v})\vec{w}\,, \tag{8.28}$$

to expand the two terms and exhibit the cancellation of the terms $(\vec{n}\cdot\vec{v})\vec{w}$, leaving the two terms equal to $-\vec{n}\times(\vec{v}\times\vec{w})$. □

Remark that the identity for vector products in the second line of (8.27) could also be written,

$$\vec{n}\times(\vec{v}\times\vec{w}) + \vec{v}\times(\vec{w}\times\vec{n}) + \vec{w}\times(\vec{n}\times\vec{v}) = 0\,. \tag{8.29}$$

This is a "Jacobi-type" identity for triple vector products.
We claim that

$$\boxed{\sum_{m=1}^{3}\epsilon_{i\ell m}\epsilon_{mjk} = \delta_{ij}\delta_{\ell k} - \delta_{ik}\delta_{\ell j}}\,, \tag{8.30}$$

which is equivalent to an identity for any four 3-vectors $\vec{A},\vec{B},\vec{C},\vec{D}$, namely

$$\boxed{(\vec{A}\times\vec{B})\cdot(\vec{C}\times\vec{D}) = \left(\vec{A}\cdot\vec{C}\right)\left(\vec{B}\cdot\vec{D}\right) - \left(\vec{A}\cdot\vec{D}\right)\left(\vec{B}\cdot\vec{C}\right)}\,. \tag{8.31}$$

One sees (8.31) using the volume relation $(\vec{B}\times\vec{A})\cdot\vec{E} = \vec{B}\cdot(\vec{A}\times\vec{E})$, along with the triple scalar product formula, $\vec{A}\times(\vec{C}\times\vec{D}) = (\vec{A}\cdot\vec{D})\,\vec{C} - (\vec{A}\cdot\vec{C})\,\vec{D}$. Thus

$$\begin{aligned}(\vec{A}\times\vec{B})\cdot(\vec{C}\times\vec{D}) &= -(\vec{B}\times\vec{A})\cdot(\vec{C}\times\vec{D}) = -\vec{B}\cdot(\vec{A}\times(\vec{C}\times\vec{D})) \\ &= -\vec{B}\cdot\Big((\vec{A}\cdot\vec{D})\,\vec{C}) - (\vec{A}\cdot\vec{C})\,\vec{D}\Big)\,. \end{aligned}\tag{8.32}$$

We end this section with a useful identity.

**Proposition 8.3.** *For any 3-vector $\vec{v}$ under rotations, we have*

$$\left[(\vec{p}\times\vec{L})_i\,,\,v_j\right]_\xi = (\vec{L}\times\nabla_{\vec{x}})_i v_j + (\vec{p}\cdot\vec{v})\,\delta_{ij} - v_i\,p_j\,, \tag{8.33}$$

*which entails*

$$\left[(\vec{p}\times\vec{L})_i\,,\,v_j\right]_\xi - \left[(\vec{p}\times\vec{L})_j\,,\,v_i\right]_\xi \;=\; \sum_{k=1}^{3}\epsilon_{ijk}\left(\left(\vec{L}\times\nabla_{\vec{x}}\right)\times\vec{v}+\vec{p}\times\vec{v}\right)_k \;. \tag{8.34}$$

*Proof.* Using the product rule for Poisson brackets,

$$\begin{aligned}\left[(\vec{p}\times\vec{L})_i\,,\,v_j\right]_\xi &= \sum_{\ell,m=1}^{3}\epsilon_{i\ell m}\,[p_\ell L_m\,,\,v_j]_\xi = \sum_{\ell,m=1}^{3}\epsilon_{i\ell m}\left([p_\ell\,,\,v_j]_\xi\,L_m + p_\ell\,[L_m\,,\,v_j]_\xi\right)\\ &= \sum_{\ell,m=1}^{3}\epsilon_{i\ell m}\left(-\frac{\partial v_j}{\partial x_\ell}L_m + p_\ell\sum_{k=1}^{3}\epsilon_{mjk}\,v_k\right)\;.\end{aligned} \tag{8.35}$$

Using (8.30) one can rewrite (8.35) as

$$\left[(\vec{p}\times\vec{L})_i\,,\,v_j\right]_\xi \;=\; (\vec{L}\times\nabla_{\vec{x}})_i v_j + \sum_{\ell,k=1}^{3}\left(\delta_{ij}\delta_{\ell k}-\delta_{ik}\delta_{\ell j}\right)p_\ell v_k\;, \tag{8.36}$$

which is (8.33). The relation (8.34) follows immediately, □

8.4. **The group $SO(4)$ of rotations in $\mathbb{R}^4$.** The group $SO(4)$ is the group of $4\times 4$ real orthogonal matrices with determinant $+1$. We can represent this group as acting on the configuration space $\mathbb{R}^4$ with coordinates $x=(x_1,x_2,x_3,x_4)$, and real orthogonal matrices leave invariant the Euclidean length of $x$, as $\langle Rx,Rx\rangle = \langle x, R^{\mathrm{tr}}Rx\rangle = \langle x,x\rangle$ for all $x$.

One can also consider a representation of $SO(4)$ on an 8-dimensional phase space with coordinates $p=(p_1,p_2,p_3,p_4)$ and $\xi=(x,p)=(x_1,x_2,x_3,x_4,p_1,p_2,p_3,p_4)$. On this phase space $\mathbb{R}^8$ one can define 6 independent generators

$$L_{ij} = x_i p_j - x_j p_i\;, \tag{8.37}$$

for rotations in the $(ij)$ plane of $x$ and of $p$. Here $i,j=1,2,3,4$. These are like components of the angular momentum in three space, but we can make such a definition in any dimension.

It is interesting to write down the corresponding derivatives $D_{L_{ij}}$ that act on phase space. They are

$$-\,D_{L_{ij}} = \left(x_i\,\frac{\partial}{\partial x_j} - x_j\,\frac{\partial}{\partial x_i}\right) + \left(p_i\,\frac{\partial}{\partial p_j} - p_j\,\frac{\partial}{\partial p_i}\right)\;. \tag{8.38}$$

Clearly these are the generators that specify infinitesimal rotations in the $(x_i,x_j)$-plane as well as in the $(p_i,p_j)$-plane. The derivative $D_{L_{ij}}$ gives zero if one differentiates any function on phase space not depending on $x_i,x_j,p_i$, or $p_j$. Furthermore

$$-\,D_{L_{ij}}\left(x_i^2+x_j^2\right) = -D_{L_{ij}}\left(p_i^2+p_j^2\right) = 0\;. \tag{8.39}$$

Thus the scalars

$$x^2 = x_1^2+x_2^2+x_3^3+x_4^2\;,\quad\text{and}\quad p^2 = p_1^2+p_2^2+p_3^3+p_4^2\;, \tag{8.40}$$

satisfy $D_{L_{ij}}x^2 = D_{L_{ij}}p^2 = 0$, for all $i, j$. Thus, as expected, we infer that the canonical transformations $e^{-\lambda_{ij}\,D_{L_{ij}}}$ leave the squared Euclidean lengths $x^2$ of $x$ and $p^2$ of $p$ invariant,

$$e^{-\lambda_{ij}\,D_{L_{ij}}}x^2 = x^2 \ , \quad \text{and} \quad e^{-\lambda_{ij}\,D_{L_{ij}}}p^2 = p^2 \ . \tag{8.41}$$

This shows that for 6 real parameters $\lambda_{ij} = -\lambda_{ji}$, which we denote collectively by $\lambda$, the 6-parameter group of transformations of the form

$$T(\lambda) = e^{-\frac{1}{2}\sum_{i,j=1}^{4}\lambda_{ij}\,D_{L_{ij}}} \tag{8.42}$$

is a representation of the group $SO(4)$. Restricted to act on the coordinates $\xi$ on phase space,

$$T(\lambda)\xi = \begin{pmatrix} R(\lambda) & 0 \\ 0 & R(\lambda) \end{pmatrix} \xi \ , \quad \text{where} \quad R(\lambda) \in SO(4) \ . \tag{8.43}$$

There is a very simple form to collect the generators $L_{ij}$ which is special to dimension 4. Arrange them in a skew-symmetric matrix, as

$$L_{ij} = \begin{pmatrix} 0 & L_3 & -L_2 & K_1 \\ -L_3 & 0 & L_1 & K_2 \\ L_2 & -L_1 & 0 & K_3 \\ -K_1 & -K_2 & -K_3 & 0 \end{pmatrix}_{ij} \ . \tag{8.44}$$

This arrangement shows that there are two natural sets of 3 components, $L_i$ and $K_i$. They are

$$L_i = \frac{1}{2}\sum_{j,k=1}^{3}\epsilon_{ijk}\,L_{jk} \ , \ \text{for} \ i = 1,2,3 \ , \quad \text{and} \quad K_i = L_{i4} \ , \ \text{for} \ i = 1,2,3 \ . \tag{8.45}$$

This leads to the Poisson bracket relations

$$[L_i \,,\, L_j]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}L_k \ , \quad [L_i \,,\, K_j]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}K_k \ , \quad [K_i \,,\, K_j]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}L_k \ . \tag{8.46}$$

Here the Poisson brackets of the $L_i$'s are the familiar Poisson brackets for the angular momentum of a single particle. We have already encountered these generators on phase space as components of angular momentum, see §8.2.1; we saw in §8.2 that they generate rotations on phase space for the coordinates $(x_1, x_2, x_3, p_1, p_2, p_3)$. So here we need only understand the $K_i$ and their relation to the $L_i$. We see, among other things, that the Poisson brackets (8.46) ensure that both $L_i$ and $K_i$ transform as three-vectors under transformations generated by the three $L_i$'s. Thus we designate them as 3-vectors $\vec{L}$ and $\vec{K}$.

In fact one can decouple two "angular momenta" in the Poisson bracket relations for $SO(4)$ by defining as new generators the vectors

$$\vec{M} = \frac{1}{2}\left(\vec{L} + \vec{K}\right) \ , \quad \text{and} \quad \vec{N} = \frac{1}{2}\left(\vec{L} - \vec{K}\right) \ . \tag{8.47}$$

The Poisson bracket relations (8.46) then yield the Poisson brackets for $\vec{M}$ and $\vec{N}$ as *independent* angular momenta 3-vectors. We find that

$$[M_i\,,\,M_j]_\xi = \sum_{k=1}^{3} \epsilon_{ijk}\,M_k\;,\quad [N_i\,,\,N_j]_\xi = \sum_{k=1}^{3} \epsilon_{ijk}\,N_k\;,\quad [M_i\,,\,N_j]_\xi = 0\;. \tag{8.48}$$

The statement that $\vec{M}$ and $\vec{N}$ are "independent" means that the Poisson brackets between any component of $\vec{M}$ and any component of $\vec{N}$ vanishes.

8.5. **The group $SO(1,3)_+$ of restricted Lorentz transformations on $R^4$.** The group $SO(1,3)_+$ is the group of real $4\times 4$ matrices acting on coordinates $x\in\mathbb{R}^4$ that have determinant $+1$, and that leave invariant the Minkowski form $x_M^2 = x_4^2 - x_1^2 - x_2^2 - x_3^2$, and that do not change the sign of $x_4$. It is sometimes called the restricted Lorentz group; it is also called the proper, orthochronous Lorentz group. See the notes on Lorentz transformations and the group $SL(2,\mathbb{C})$.

We can also analyze a representation of this group on a phase space $\xi = (x,p)\in\mathbb{R}^8$. In order to analyze this representation, define a skew-symmetric matrix of generators given by 6 independent functions on phase space. It has a form similar to the matrix we used to summarize the generators of $SO(4)$, but some of the entries are defined differently. Consider

$$L_{ij} = \begin{pmatrix} 0 & L_3 & -L_2 & K_1 \\ -L_3 & 0 & L_1 & K_2 \\ L_2 & -L_1 & 0 & K_3 \\ -K_1 & -K_2 & -K_3 & 0 \end{pmatrix}_{ij}\;, \tag{8.49}$$

with

$$L_i = \frac{1}{2}\sum_{j,k=1}^{3} \epsilon_{ijk}\,L_{jk}\;,\ \text{ for }\ i=1,2,3\;,\quad \text{and}\quad K_i = L_{i4}\;,\ \text{ for }\ i=1,2,3\;. \tag{8.50}$$

We keep the definition of the three generators $L_i$ that we introduced in the study of $SO(4)$, but we give a new definition for the three generators $K_i$. Thus as in §8.4,

$$L_i = \frac{1}{2}\sum_{j,k=1}^{3} \epsilon_{ijk}\,(x_j p_k - x_k p_j)\;,\quad \text{for}\quad i=1,2,3\;. \tag{8.51}$$

On the other hand, for the $K_i$ we substitute a $+$ sign for the $-$ sign and define

$$K_i = x_i p_4 + x_4 p_i\;,\quad \text{for}\quad i=1,2,3\;. \tag{8.52}$$

With this definition, we find that for $i=1,2,3$, as with the generators of $SO(4)$,

$$-\,D_{L_i} = \frac{1}{2}\sum_{j,k=1}^{3} \epsilon_{ijk}\left(\left(x_j\,\frac{\partial}{\partial x_k} - x_k\,\frac{\partial}{\partial x_j}\right) + \left(p_j\,\frac{\partial}{\partial p_k} - p_k\,\frac{\partial}{\partial p_j}\right)\right)\;. \tag{8.53}$$

But now

$$-D_{K_i} = \left(x_i\frac{\partial}{\partial x_4}+x_4\frac{\partial}{\partial x_i}\right)-\left(p_i\frac{\partial}{\partial p_4}+p_4\frac{\partial}{\partial p_i}\right) . \tag{8.54}$$

Introduce the Minkowski-space scalars,

$$x_M^2 = x_4^2 - x_1^2 - x_2^2 - x_3^3 , \quad \text{and} \quad p_M^2 = p_4^2 - p_1^2 - p_2^2 - p_3^3 . \tag{8.55}$$

Similar to our argument above in the case of $SO(4)$, with our new definitions (8.54) one has

$$D_{L_i}x_M^2 = D_{K_i}x_M^2 = D_{L_i}p_M^2 = D_{K_i}p_M^2 = 0 . \tag{8.56}$$

Thus the transformations generated by $D_{L_i}$ and $D_{K_i}$ preserve the square of the Minkowski length of $x$ and of $p$. Given real parameters $\vec{\lambda}, \vec{\mu}$ with components $\lambda_i, \mu_i$ for $i = 1, 2, 3$, one has

$$e^{-\mathcal{D}_{\vec{\lambda}\cdot\vec{L}+\vec{\mu}\cdot\vec{K}}}\, x_M^2 = x_M^2 , \quad \text{and} \quad e^{-\mathcal{D}_{\vec{\lambda}\cdot\vec{L}+\vec{\mu}\cdot\vec{K}}}\, p_M^2 = p_M^2 . \tag{8.57}$$

The corresponding Poisson brackets for the generators of this group of transformations satisfy

$$[L_i\,,\,L_j]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}L_k\,, \quad [L_i\,,\,K_j]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}K_k\,, \quad [K_i\,,\,K_j]_\xi = -\sum_{k=1}^{3}\epsilon_{ijk}L_k\,. \tag{8.58}$$

These relations are very similar to the Poisson bracket relations (8.46) for the generators of the group $SO(4)$, except for the minus sign in the third set of PB's. In fact the transformations $e^{-\mathcal{D}_{\vec{\lambda}\cdot\vec{L}+\vec{\mu}\cdot\vec{K}}}$ are restricted Lorentz transformations. (See the notes on Lorentz transformations and $SL(2,\mathbb{C})$.)

In order to identify how the 6 parameters $\vec{\lambda}$ and $\vec{\mu}$ match with the usual parameters in the Lorentz group, it is of interest to check what explicit transformation $K_i$ generates on phase space. In particular, we obtain a physical interpretation by computing what is $D_{K_i}$. This lets us determine the action of $e^{-\chi D_{K_i}}$ on the phase space coordinates.

Thus we see that on configuration space ($x$-space) the transformation $D_{K_i}$, for $i = 1, 2$, or 3, affects only the coordinates $x_i$ and $x_4$. Likewise on momentum space ($p$-space) the transformation $D_{K_i}$ affects only the coordinates $p_i$ and $p_4$. And for a fixed unit vector $\vec{n}$ in three-space, $D_{\vec{K}\cdot\vec{n}}$ affects only the components of $x$ in direction $\vec{n}$ or the component $x_4$. Likewise $D_{\vec{K}\cdot\vec{n}}$ affects the same components of $p$.

Let us compute the transformation $e^{-\chi D_{K_3}}$ on phase space. This transformation acts only on the components $x_3, x_4, p_3, p_4$. So we can ignore the other 4 coordinate directions and represent the action of the transformation $e^{-\chi D_{K_3}}$ as a $4\times 4$ block-diagonal matrix made up of $2\times 2$ blocks $\Lambda_1, \Lambda_2$. Here $\Lambda_1 = \Lambda_1(\chi)$ transforms $(x_3, x_4)$ among themselves, and $\Lambda_2 = \Lambda_2(\chi)$ transforms $(p_3, p_4)$ among themselves. Furthermore, it is clear for the form of (8.54) that $\Lambda_1 = \Lambda_2^{-1} = \Lambda(\chi)$. Therefore if we look at the subspace of phase space with coordinates $\eta = (x_3, x_4, p_3, p_4)$, we find that the transformation $e^{-\chi D_{K_3}}$ has the form of a block-diagonal matrix

$$e^{-\chi D_{K_3}}\,\eta = \begin{pmatrix}\Lambda(\chi) & 0\\ 0 & \Lambda(\chi)^{-1}\end{pmatrix}\eta\,. \tag{8.59}$$

So to determine the transformation $e^{-\chi D_{K_3}}$, we need only find the $2 \times 2$ matrix $\Lambda(\chi)$ that transforms $(x_3, x_4)$.

We now find $\Lambda = \Lambda(\chi)$. Note that

$$-D_{K_3} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_4 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} . \tag{8.60}$$

Hence

$$(-D_{K_3})^{2n} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} , \quad \text{and} \quad (-D_{K_3})^{2n+1} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} , \tag{8.61}$$

so

$$\boxed{e^{-\chi D_{K_3}} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \cosh\chi & \sinh\chi \\ \sinh\chi & \cosh\chi \end{pmatrix} \begin{pmatrix} x_3 \\ x_4 \end{pmatrix}} , \quad \text{so} \quad \boxed{\Lambda = \begin{pmatrix} \cosh\chi & \sinh\chi \\ \sinh\chi & \cosh\chi \end{pmatrix}} . \tag{8.62}$$

This justifies our calling the group generated by the functions $\vec{L}$ and $\vec{K}$ the proper Lorentz group. This group leaves invariant the quadratic form $x_1^2 + x_2^2 + x_3^2 - x_4^2$, so it is also called the group $SO(3,1)_+$.

This matrix $\Lambda$ is the matrix that describes a Lorentz boost by velocity $v = c \tanh\chi$ in the direction $x_3$. Acting on the phase space coordinates $x_3, x_4, p_3, p_4$ the matrix

$$e^{-\chi D_{K_3}} \eta = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda^{-1} \end{pmatrix} \eta . \tag{8.63}$$

Note that this transformation has the form (4.8) of the Jacobian of a canonical transformation arising from a generator $G = \langle p, Tq \rangle$ for a symmetric matrix $T$, see (5.21). In this case $T_{k\ell} = \delta_{k4}\,\delta_{\ell i} + \delta_{i4}\,\delta_{\ell k}$.

8.6. **Symmetry of the Kepler Problem.** Here we find the symmetry groups connected with the two-body Kepler problem in relative coordinates. Ignore the center of mass motion, and consider the Hamiltonian

$$H = \frac{\vec{p}^{\,2}}{2\mu} - \frac{k}{r} . \tag{8.64}$$

We denote the relative coordinate to the force center by $\vec{x}$, with $r = |\vec{x}|$. We show that for negative energies (bound orbits) there is an $SO(4)$ on the phase space. In the case of positive energies (hyperbolic scattering orbits) the symmetry is described by the group of proper Lorentz transformations that do not reverse the sign of time. We call this group $SO(1,3)_+$ in §8.5.

Let us focus on the two vectors, the angular momentum and the Lenz vector,

$$\vec{L} = \vec{x} \times \vec{p} , \qquad \vec{\epsilon} = \frac{\vec{p} \times \vec{L}}{\mu k} - \vec{n} , \quad \text{where } \vec{n} = \frac{\vec{x}}{r} . \tag{8.65}$$

We have shown both $\vec{L}$ and $\vec{\epsilon}$ are conserved by $H$. Thus $\left[ H \,,\, \vec{L} \right]_\xi = 0$ and $[H \,,\, \vec{\epsilon}\,]_\xi = 0$.

**Theorem 8.4.** *The Poisson bracket relations between the components of $\vec{L}$ and $\vec{\epsilon}$ are*

$$[L_i\,,\,L_j]_\xi = \sum_{k=1}^{3} \epsilon_{ijk} L_k\;,\quad [L_i\,,\,\epsilon_j]_\xi = \sum_{k=1}^{3} \epsilon_{ijk}\epsilon_k\;,\quad [\epsilon_i\,,\,\epsilon_j]_\xi = -\frac{2H}{\mu k^2}\sum_{k=1}^{3}\epsilon_{ijk} L_k\;. \tag{8.66}$$

*Proof.* We already have verified the relations for $[L_1\,,\,L_j]_\xi$. The relations for $[L_i\,,\,\epsilon_j]_\xi$ state that $\vec{\epsilon}$ is a 3-vector under rotations. This is a consequence of the fact the $\vec{p}$ and $\vec{L}$ are both 3-vectors, so we infer from Proposition 8.2 that $\vec{p}\times\vec{L}$ is also a 3-vector. Furthermore Proposition 8.2 ensures that $\vec{n}$ is a 3-vector, because it is the product of the 3-vector $\vec{x}$ and the scalar $r$. The sum of 3-vectors is a 3-vector, so we conclude that $\vec{\epsilon}$ is a 3-vector.

This means that we need only establish the relations $[\epsilon_i\,,\,\epsilon_j]_\xi$ , which requires a bit more computation. It is sufficient to verfiy that $[\epsilon_1\,,\,\epsilon_2]_\xi = -(2H/\mu k^2)L_3$. The remaining cases follow from by rotation of this relation and skew-symmetry of the Poisson bracket. So let us inspect

$$\begin{aligned}(\mu k)^2\,[\epsilon_1\,,\,\epsilon_2]_\xi &= \left[(\vec{p}\times\vec{L})_1 - \mu k\frac{x_1}{r}\,,\,(\vec{p}\times\vec{L})_2 - \mu k\frac{x_2}{r}\right]_\xi \\ &= \left[(\vec{p}\times\vec{L})_1\,,\,(\vec{p}\times\vec{L})_2\right]_\xi + \left[(\vec{p}\times\vec{L})_1\,,\,-\mu k\frac{x_2}{r}\right]_\xi - \left[(\vec{p}\times\vec{L})_2\,,\,-\mu k\frac{x_1}{r}\right]_\xi\;.\end{aligned} \tag{8.67}$$

For the first of the three Poisson brackets on the right-hand side, we use a special case of Proposition 8.3, coming from the choice: $\vec{v}=\vec{p}\times\vec{L}$, $i=1$, and $j=2$. In this case,

$$v_2 \;=\; (\vec{p}\times\vec{L})_2 = -(x_1p_1+x_3p_3)p_2 + x_2(p_1^2+p_3^2)\;. \tag{8.68}$$

Putting (8.68) into (8.33) gives the desired result,

$$\begin{aligned}\left[(\vec{p}\times\vec{L})_1\,,\,(\vec{p}\times\vec{L})_2\right]_\xi &= -\left(p_1^2+p_3^2\right)L_3 - p_2p_3L_2 - p_2(p_2L_3 - p_3L_2) \\ &= -\left(p_1^2+p_2^2+p_3^2\right)L_3 = -\vec{p}^{\,2}L_3\;.\end{aligned} \tag{8.69}$$

For the other terms in (8.67), we use (8.34) in Proposition 8.3, with the choice $\vec{v}=-\mu k\vec{n}$ and $\vec{n}$ the unit vector with components

$$v_\ell = -\mu k\frac{x_\ell}{r}\;. \tag{8.70}$$

Then

$$(\nabla_{\vec{x}}v_j)_\ell \;=\; -\mu k\left(\delta_{\ell j}\left(\frac{1-n_\ell^2}{r}\right) - (1-\delta_{\ell j})\,\frac{n_\ell n_j}{r}\right) \quad\text{and}\quad \vec{p}\times\vec{v} = \mu k\frac{\vec{L}}{r}\;. \tag{8.71}$$

Using (8.71) and (8.34) with $i=1$ and $j=2$ gives for the last two Poisson brackets in (8.67),

$$\begin{aligned}
&\left[(\vec{p}\times\vec{L})_1\,,\,-\mu k\frac{x_2}{r}\right]_\xi-\left[(\vec{p}\times\vec{L})_2\,,\,-\mu k\frac{x_1}{r}\right]_\xi\\
&\qquad=\left(\left(\vec{L}\times\nabla_{\vec{x}}\right)\times\vec{v}+\vec{p}\times\vec{v}\right)_3\\
&\qquad=\left(\vec{L}\times\nabla_{\vec{x}}\right)_1 v_2-\left(\vec{L}\times\nabla_{\vec{x}}\right)_2 v_1+(\vec{p}\times\vec{v})_3\\
&\qquad=\left(L_2\frac{\partial}{\partial x_3}-L_3\frac{\partial}{\partial x_2}\right)v_2-\left(L_3\frac{\partial}{\partial x_1}-L_1\frac{\partial}{\partial x_3}\right)v_1+\mu k\frac{L_3}{r}\\
&\qquad=L_3\left(-\frac{\partial}{\partial x_2}v_2-\frac{\partial}{\partial x_1}v_1-\mu k\frac{n_3^2}{r}+\mu k\frac{1}{r}\right)=L_3\,\frac{2\mu k}{r}\;. \qquad (8.72)
\end{aligned}$$

In the penultimate equality in (8.72), we have eliminated the $L_1$ and $L_2$ terms using the relation $\vec{L}\cdot\vec{n}=0$, namely we use

$$\left(L_2\frac{\partial}{\partial x_3}\right)v_2+\left(L_1\frac{\partial}{\partial x_3}\right)v_1=(L_2n_2+L_1n_1)\,n_3\,\frac{\mu k}{r}=-L_3\left(\mu k\frac{n_3^2}{r}\right)\;. \qquad (8.73)$$

And in the last equality of (8.72), we find using (8.71) that the coefficient of $L_3$ coming from the four terms is

$$\mu k\left(\frac{1-n_2^2}{r}+\frac{1-n_1^2}{r}-\frac{n_3^2}{r}+\frac{1}{r}\right)=\frac{2\mu k}{r}\;. \qquad (8.74)$$

Inserting (8.69) and (8.72) into (8.67), and dividing by $(\mu k)^2$, we have shown that

$$[\epsilon_1\,,\,\epsilon_2]_\xi=\frac{1}{\mu^2k^2}\left(-\vec{p}^{\,2}+\frac{2\mu k}{r}\right)L_3=-\frac{2}{\mu k^2}\left(\frac{\vec{p}^{\,2}}{2\mu}-\frac{k}{r}\right)L_3=-\frac{2H}{\mu k^2}L_3\;, \qquad (8.75)$$

as claimed. □

In order to combine the relations between $\vec{L}$ and $\vec{\epsilon}$, we redefine the conserved Lenz vector to have the same dimension as $\vec{L}$. Multiplying by a constant or a function of $H$ does not introduce new Poisson brackets. But the result for the Poisson brackets (8.66) suggests treating separately the cases of positive energy (hyperbolic orbits) and negative energy (elliptic orbits).

8.7. **The Negative Energy Case.** In this case it is convenient to define the vector

$$\vec{K}=\left(-\frac{\mu k^2}{2H}\right)^{1/2}\,\vec{\epsilon}\,. \qquad (8.76)$$

Then one can rewrite the relations (8.66) in the form

$$\boxed{[L_i\,,\,L_j]_\xi=\sum_{k=1}^{3}\epsilon_{ijk}L_k\;,\quad [L_i\,,\,K_j]_\xi=\sum_{k=1}^{3}\epsilon_{ijk}K_k\;,\quad [K_i\,,\,K_j]_\xi=\sum_{k=1}^{3}\epsilon_{ijk}L_k}\;. \qquad (8.77)$$

These are the Lie algebra relations for the group $SO(4)$, namely the group of proper rotations in 4-space. This group has finite volume and is said to be *compact*. See §8.4. Furthermore, using the formula for the eccentricity vector derived in the notes on the Kepler problem, $\epsilon^2 = 1 + \frac{2EL^2}{\mu k^2}$, we have

$$\vec{K}^{\,2} = -\frac{\mu k^2}{2H}\vec{\epsilon}^{\,2} = -\frac{\mu k^2}{2H}\left(1 + \frac{2H\vec{L}^{\,2}}{\mu k^2}\right) = -\frac{\mu k^2}{2H} - \vec{L}^{\,2} \ , \tag{8.78}$$

and writing $E$ as $H$, one can express the Hamiltonian in as an elementary function of the two conserved vectors $\vec{L}$ and $\vec{K}$,

$$H = \frac{\mu k^2}{2\left(\vec{L}^{\,2} + \vec{K}^{\,2}\right)} \ . \tag{8.79}$$

As explained in §8.4, one can then combine the vectors $\vec{K}$ and $\vec{L}$ to form two *independent* angular momentum vectors by defining

$$\vec{M} = \frac{1}{2}\left(\vec{L} + \vec{K}\right) \ , \quad \text{and} \quad \vec{N} = \frac{1}{2}\left(\vec{L} - \vec{K}\right) \ , \tag{8.80}$$

Each of these two vectors have Poisson bracket relations similar to the angular momentum brackets, but they are decoupled—in the sense that the brackets between components of the different vectors vanish. One has

$$\boxed{\left[M_i\,,\,M_j\right]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}M_k \ , \quad \left[N_i\,,\,N_j\right]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}N_k \ , \quad \left[M_i\,,\,N_j\right]_\xi = 0} \ . \tag{8.81}$$

Since $\vec{L}\cdot\vec{\epsilon} = 0$, also $\vec{L}\cdot\vec{K} = 0$, and $\vec{M}^{\,2} = \frac{1}{4}\left(\vec{L}^{\,2} + \vec{K}^{\,2}\right) = \vec{N}^{\,2}$. Hence in terms of the independent vectors $\vec{M}$ and $\vec{N}$,

$$H = \frac{\mu k^2}{4\left(\vec{M}^2 + \vec{N}^2\right)} \ . \tag{8.82}$$

8.8. **The Positive Energy Case.** The symmetry group for positive energy (unbounded orbits) is different. In this case, it is convenient to define $\vec{K}$ *not* as (8.76), but rather as

$$\vec{K} = \left(\frac{\mu k^2}{2H}\right)^{1/2} \vec{\epsilon} \, . \tag{8.83}$$

Then in place of (8.77), one obtains the relations

$$\boxed{\left[L_i\,,\,L_j\right]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}L_k \ , \quad \left[L_i\,,\,K_j\right]_\xi = \sum_{k=1}^{3}\epsilon_{ijk}K_k \ , \quad \left[K_i\,,\,K_j\right]_\xi = -\sum_{k=1}^{3}\epsilon_{ijk}L_k} \ . \tag{8.84}$$

These are just the Poisson bracket relations (8.58) found for the group $SO(1,3)_+$ of Lorentz transformations, see §8.5. This group has infinite volume and is said to be *non-compact.* This property arisis from the fact that Lorentz boosts can have velocity less in magnitude than, but never equal in magnitude to, the velocity of light.

In this case a real combination of $\vec{L}$ and $\vec{K}$, as in (8.80), does not decouple two algebras. Now we need to introduce complex functions on phase space! The fact that we are led to complex functions $\vec{M}$ and $\vec{N}$ on phase space is associated with the fact that the Poisson bracket algebra (8.84) characterizes a symmetry for a non-compact group (actually the group of proper Lorentz transformations). On 4-dimensional space-time: a rotation is a change of coordinates given by a real, orthogonal matrix, while a Lorentz boost is given by a real, symmetric matrix.

In order to decouple $\vec{M}$ and $\vec{N}$, take the *complex* combinations

$$\vec{M} = \frac{1}{2}\left(\vec{L} + i\vec{K}\right) \ , \quad \text{and} \quad \vec{N} = \frac{1}{2}\left(\vec{L} - i\vec{K}\right) \ . \tag{8.85}$$

Then

$$[M_i\,,\,M_j]_\xi = \sum_{k=1}^{3} \epsilon_{ijk} M_k \ , \quad [N_i\,,\,N_j]_\xi = \sum_{k=1}^{3} \epsilon_{ijk} N_k \ , \quad [M_i\,,\,N_j]_\xi = 0 \ . \tag{8.86}$$

8.9. **Comment.** We have thus identified the Poisson bracket relations between $\vec{L}$ and $\vec{K}$ for negative and for positive energy, and the groups that $e^{-D_{\langle\vec{\lambda},\vec{L}\rangle+\langle\vec{\mu},\vec{K}\rangle}}$. These are groups on the phase space $\mathbb{R}^6$. However we have not found a formula on phase space for the action of the time translation group $e^{-tD_H}$. Nor do we have an explicit connection between the solution of the Hamilton equations on the phase space $\mathbb{R}^6$ and the symmetry transformations which we worked out on the phase space $\mathbb{R}^8$. This has been carried out and provides interesting reading![3] It is related to the symmetry of the hydrogen atom in quantum theory, which also has been analyzed, and is related.[4] The original use of this symmetry, in order to solve for the bound state energies of the hydrogen atom in quantum theory was given by Pauli.[5]

[3]Thomas Ligon and Manfred Schaaf, On the Global Symmetry of the Classical Kepler Problem, *Reports on Mathematical Physics*, **9** (1976), 281–300.

[4]Myron Bander and Claude Itzykson, Group Theory of the Hydrogen Atom (I) and (II), *Rev. Mod. Phys.*, **38** (1966), 330–345, and 346–358.

[5]Wolfgang Pauli Jr., Über das Wasserstoffspektrum vom Standpunkt der neuen Quantunmechanik, *Z. Phys.*, **36** (1926), 336–363. An English translation appears in *Sources of Quantum Mechanics*, Edited with a Historical Introduction by B. L. van der Waerden, pp 387–415, North Holland Publishing Company, Amsterdam 1967.