Geometric and physical interpretation of the action principle

Gabriele Carcassi, Christine A. Aidala

Physics Department, University of Michigan, Ann Arbor, MI 48109

(Dated: August 15, 2023)

We give a geometric interpretation for the principle of stationary action in classical Lagrangian particle mechanics. In a nutshell, the difference of the action along a path and its variation effectively “counts” the possible evolutions that “go through” the area enclosed. If the path corresponds to a possible evolution, all neighboring evolutions will be parallel, making them tangent to the area enclosed by the path and its variation, thus yielding a stationary action. This treatment gives a full physical account of the geometry of both Hamiltonian and Lagrangian mechanics which is founded on three assumptions: determinism and reversible evolution, independence of the degrees of freedom and equivalence between kinematics and dynamics. The logical equivalence between the three assumptions and the principle of stationary action leads to a much cleaner conceptual understanding.

I. INTRODUCTION

While the principle of stationary action is regarded by many as one of the most important tools in physics, its physical meaning is not completely clear.[1–5] First of all, the typical characterization of the Lagrangian as the difference between kinetic and potential energy fails even for simple systems, like a charged particle under a magnetic field. Moreover, the Lagrangian for a system is not uniquely defined, making the actual value of the action for a path not directly physically significant. We are left to wonder: what exactly is the action and why is it stationary for actual trajectories?

As part of our larger project Assumptions of Physics, we developed an approach, called Reverse Physics [6], which examines current theories to find a set of starting physical assumptions that are sufficient to rederive them. We have found that Lagrangian mechanics is equivalent to three assumptions: determinism/reversibility, independence of degrees of freedom and kinematics/dynamics equivalence [7]. This physically motivated understanding of the classical theory can be used to characterize both the physics and the geometry underlying the principle of stationary action. What we find is that this arises as a general mathematical feature of divergence-free fields (and closed two-forms), which are the appropriate tools to describe a flow that conserves the number of states. The assumption of equivalence between kinematics and dynamics (i.e. we can reconstruct the dynamical state simply by looking at the trajectory) is what allows us to express the principle in the usual form. The argument can proceed in the reverse direction: assuming the principle of stationary action recovers a dynamical system that exhibits those three physical assumptions.

The mathematics needed to run the argument is well established[8–11]. The two aspects that have been sorely lacking, and that we provide, are (1) a clear geometric interpretation of the action principle and (2) a tight connection between the math and the physics it represents. [This has been the hardest problem to solve, as the modern mathematical concepts and notation have departed from physical intuition. We will therefore use, instead of the standard notation of differential geometry, an extended version of the one physicists use in general relativity.] To make the result accessible to the widest audience, we first cover the case of a single degree of freedom using standard vector calculus. We then proceed to the case of multiple independent degrees of freedom using tools from differential geometry. The main article will present all the key points needed to follow the argument and its physical and geometric meaning, leaving the mathematical details and calculations to the Supplementary Information.

II. ONE DEGREE OF FREEDOM

As we want to characterize the evolution of states over time, the appropriate setting is phase space extended with the time variable[12, 13]. That is, the space charted by position q, momentum p and time t as can be seen in Fig. 1. In the same way that we write $x^i = [x, y, z]$ for the three dimensions of space, we write

$$\xi^a = [q, p, t]$$  \hspace{1cm} (1)

for the three dimensions of the extended phase space.

Under the assumption that

the system undergoes deterministic and reversible evolution  \hspace{1cm} (DR)

FIG. 1. Evolutions in the extended phase space and the divergence-free displacement field.
we can define a displacement vector field
\[
\vec{S} = \begin{bmatrix} dq \quad dp \quad dt \end{bmatrix} = S^a e_a.
\] (2)

that describes how states move in time.\(^1\) In dynamical system literature, this is referred to as the vector field of the dynamical system. The time component of the displacement vector field is constrained, as we have
\[
S^t = \frac{dt}{dt} = 1.
\] (3)

If assumption (DR) is valid, we expect the flow of states through a closed surface to be zero: as many states flow in as flow out of the region. Alternatively, if we assign a probability, or probability density, to each trajectory, the assumption requires that probability not to change, so integrating the probability over a closed surface must yield zero. However we see it, assumption (DR) means the field is divergence-free.\(^2\) That is,
\[
\nabla \cdot \vec{S} = \partial_a S^a = 0.
\] (4)

Since the displacement field is divergence-free, it admits a vector potential. We have
\[
\vec{\theta} = [\theta_q, \theta_p, \theta_t] = \theta_a e^a
\]
\[
\vec{S} = -\nabla \times \vec{\theta} = -e^{abc} \partial_b \theta_c e_a.
\] (5)
The minus sign is introduced to match conventions. Mathematically, this is analogous to what is done for a magnetic field or for an incompressible fluid.

Because the displacement field must satisfy 3, without loss of generality we can set
\[
\vec{\theta} = [p, 0, -H(q, p, t)] = p e^q - H(q, p, t) e^t,
\] (6)
where \(H\) is a suitable function of \(q, p\) and \(t\). The potential \(\vec{\theta}\) is closely related to the canonical one-form of symplectic geometry and the contact form of contact geometry. By applying definition 2 and expanding 5 with 6, we have
\[
\begin{bmatrix} dq \quad dp \quad dt \end{bmatrix} = -\nabla \times \vec{\theta} = \begin{bmatrix} \frac{\partial H}{\partial p} - \frac{\partial H}{\partial q} \end{bmatrix},
\] (7)
which yields Hamilton’s equations. Note that the argument works in reverse: any Hamiltonian system with one degree of freedom yields a divergence-free displacement field, and therefore satisfies (DR).

\(^1\) Where possible, we will be writing the same expression in both vector calculus and component notations.

\(^2\) Given that this is a three-dimensional space, we can use the standard tools of vector calculus.

\[\begin{align*}
-S^t &= \frac{dt}{dt} = 1. \\
\nabla \cdot \vec{S} &= \partial_a S^a = 0. \\
\vec{\theta} &= [\theta_q, \theta_p, \theta_t] = \theta_a e^a \\
\vec{S} &= -\nabla \times \vec{\theta} = -e^{abc} \partial_b \theta_c e_a.
\end{align*}\] (5)

The minus sign is introduced to match conventions. Mathematically, this is analogous to what is done for a magnetic field or for an incompressible fluid.

Because the displacement field must satisfy 3, without loss of generality we can set
\[\vec{\theta} = [p, 0, -H(q, p, t)] = p e^q - H(q, p, t) e^t,\] (6)
where \(H\) is a suitable function of \(q, p\) and \(t\). The potential \(\vec{\theta}\) is closely related to the canonical one-form of symplectic geometry and the contact form of contact geometry. By applying definition 2 and expanding 5 with 6, we have
\[\begin{bmatrix} dq \quad dp \quad dt \end{bmatrix} = -\nabla \times \vec{\theta} = \begin{bmatrix} \frac{\partial H}{\partial p} - \frac{\partial H}{\partial q} \end{bmatrix},\] (7)
which yields Hamilton’s equations. Note that the argument works in reverse: any Hamiltonian system with one degree of freedom yields a divergence-free displacement field, and therefore satisfies (DR).

\[\begin{align*}
-S^t &= \frac{dt}{dt} = 1. \\
\nabla \cdot \vec{S} &= \partial_a S^a = 0. \\
\vec{\theta} &= [\theta_q, \theta_p, \theta_t] = \theta_a e^a \\
\vec{S} &= -\nabla \times \vec{\theta} = -e^{abc} \partial_b \theta_c e_a.
\end{align*}\] (5)

The minus sign is introduced to match conventions. Mathematically, this is analogous to what is done for a magnetic field or for an incompressible fluid.

Because the displacement field must satisfy 3, without loss of generality we can set
\[\vec{\theta} = [p, 0, -H(q, p, t)] = p e^q - H(q, p, t) e^t,\] (6)
where \(H\) is a suitable function of \(q, p\) and \(t\). The potential \(\vec{\theta}\) is closely related to the canonical one-form of symplectic geometry and the contact form of contact geometry. By applying definition 2 and expanding 5 with 6, we have
\[\begin{bmatrix} dq \quad dp \quad dt \end{bmatrix} = -\nabla \times \vec{\theta} = \begin{bmatrix} \frac{\partial H}{\partial p} - \frac{\partial H}{\partial q} \end{bmatrix},\] (7)
which yields Hamilton’s equations. Note that the argument works in reverse: any Hamiltonian system with one degree of freedom yields a divergence-free displacement field, and therefore satisfies (DR).

1. Where possible, we will be writing the same expression in both vector calculus and component notations.
2. Given that this is a three-dimensional space, we can use the standard tools of vector calculus.
The last step is to express the principle exclusively in terms of kinematic variables: position, time and velocity. This can be done if we assume that

**the kinematics of the system is enough to reconstruct its dynamics.** (KE)

This means that by looking at just the trajectory in space $q(t)$, we are able to reconstruct the state at each moment in time. Therefore we must be able to write $p = p(q, \dot{q})$, and therefore we can also write

$$\delta \int_{t_1}^{t_2} \dot{q} \cdot d\gamma = \delta \int_{t_1}^{t_2} \frac{dp}{dt} \cdot d\gamma = \delta \int_{t_1}^{t_2} \left( p \frac{dq}{dt} - H \right) dt$$

$$= \delta \int_{t_1}^{t_2} L(q, \dot{q}, t) dt = 0. \quad (10)$$

We find that a system for which (DR) and (KE) are valid can be characterized in terms of the principle of stationary action with a suitable Lagrangian. The converse is also true: if the principle of stationary action allows for a unique solution, then the conjugate momentum and Hamiltonian are well defined and the system satisfies both (DR) and (KE).

We have thus demystified the principle of stationary action, and turned it into a geometric property: requiring the principle of stationary action is equivalent to requiring that the solutions are the field lines of a divergence-free field in phase space. We also have a clear physical meaning: the principle of stationary action is equivalent to assuming determinism/reversibility (DR) and kinematic equivalence (KE). However, we do feel that the principle expresses these requirements in a very roundabout way.

### III. MULTIPLE DEGREES OF FREEDOM

To generalize to multiple degrees of freedom we have to abandon the tools provided by vector calculus and embrace the ones provided by differential geometry. For $N$ independent degrees of freedom, we will have the $2N+1$ manifold charted by the variables

$$\xi^a = [q^i, p_i, t]. \quad (11)$$

The displacement vector field will be

$$\hat{S} = S^a e_a = \frac{d\xi^a}{dt} e_a = \frac{dq^i}{dt} e_q + \frac{dp_i}{dt} e_p + \frac{dt}{dt} e_t \quad (12)$$

which also satisfies 3.

The flow of $\hat{S}$ will still be divergence-free, but this property only tells us that the total number of states is conserved. We also need to assume that

**the degrees of freedom are independent.** (IND)

This means that, at each moment in time, the total number of states (i.e. the volume) of a parallelepiped will be the product of states identified by each degree of freedom (i.e. the areas). To be able to quantify the number of states identified by each degree of freedom, we introduce the rank-2 tensor (more precisely a two-form)

$$\omega = \omega_{ab} e^a \otimes e^b, \quad (13)$$

which we call the state counting form. 3

The counting form must be of rank two as independent degrees of freedom are bi-dimensional (i.e. quantity plus conjugate). Given two vectors $\hat{v}$ and $\hat{w}$, these identify a parallelogram in phase space and

$$\omega(\hat{v}, \hat{w}) = \omega_{ab} v^a w^b \quad (14)$$

quantifies the number of states over its surface.

The form $\omega$ will need to be anti-symmetric

$$\omega(\hat{v}, \hat{w}) = -\omega(\hat{w}, \hat{v}) \quad (15)$$

as the parallelogram identified by $\hat{v}$ and $\hat{w}$ in that order will be the same as the one identified by $\hat{w}$ and $\hat{v}$ with opposite orientation. The form $\omega$ will also be closed, meaning

$$\iiint_{\Sigma} \omega(d\Sigma) = 0 \quad (16)$$

over all closed contractible surfaces. [In our notation, $d\Sigma$ represents the infinitesimal surface element of integration, which is the argument of the form.] This stems from (IND). Suppose you take a surface and translate it along another independent variable. The count of states cannot depend on the independent variable. Therefore if we construct any parallelepiped at equal time, two opposite sides will contain the same number of states with opposite orientation, and therefore the integral of $\omega$ over the whole surface will be zero. Those familiar with differential geometry will recognize $\omega$ as the symplectic form of symplectic manifolds and of the symplectic bundles of contact manifolds.

Since $\omega$ is closed, in every contractible region it can be expressed as the exterior derivative of a covector $\theta$:

$$\theta = \theta_a e^a = \theta_q e^q + \theta_p e^p + \theta_t e^t$$

$$\omega = -\partial \wedge \theta = -(\partial_q \theta_p - \partial_p \theta_q) e^a \otimes e^b \quad (17)$$

The minus sign is required to match conventions. 4

The above is the generalization of 5 and it is important to note the differences. The exterior derivative takes the place of the curl and it is $\omega$, not $\hat{S}$, that can be expressed

---

3 Note that this approach is consistent with statistical mechanics[14].

4 We write the exterior derivative as $\partial \wedge \omega$ instead of $d\omega$ so that $d$ is only used for infinitesimal line or surface elements.
in terms of the potential $\theta$. Yet, there is an important geometric relationship we still need to take into account: the direction of the displacement does not contribute new states. Physically, this is a consequence of (DR): states cannot be created or disappear over time. We have

$$\omega(S, \gamma) = 0.$$  \hfill (18)

Mathematically we say that the displacement vector kills the form; it makes it zero no matter what vector we put on the other side.

Similarly to the single d.o.f., without loss of generality we can set

$$\theta = \left[p_i, 0, -H\right] = p_i e^{q^i} - H e^t.$$  \hfill (19)

and calculate the components of the counting form

$$\omega_{ab} = (-\partial \wedge \theta)_{ab} = \begin{bmatrix} 0 & \delta_i^j & \partial_q^i H \\ -\delta_j^i & 0 & \partial_p^j H \\ -\partial_q^i H & -\partial_p^j H & 0 \end{bmatrix}. \hfill (20)$$

By combining 12, 18 and 20 one finds

$$S^q_i = \partial_p^i H$$
$$S^p_i = -\partial_q^i H$$  \hfill (21)

which recovers Hamilton’s equations.

The setup to recover the principle of stationary action is the same, except that we need to substitute vector calculus with differential geometry. First we find that the flow through $\Sigma$ is the variation of the action using Stokes’ theorem

$$-\int_{\Sigma} \omega(d\Sigma) = \delta \int_{\gamma} \theta(d\gamma)$$  \hfill (22)

If $\gamma$ is a solution of the equations of motion, then $d\gamma = \dot{S} dt$. The infinitesimal surface $d\Sigma$ will be a parallelogram formed by $d\gamma$ and the direction of the variation $d\lambda$. By 18, $\dot{S}$ is the only vector field that kills the form $\omega$ therefore the solutions are such that

$$0 = \delta \int_{\gamma} \theta(d\gamma) = -\int_{\Sigma} \omega(\dot{S}, d\lambda) dt.$$  \hfill (23)

Adding assumption (KE), we can express the variational principle in the usual form

$$\delta \int_{\gamma} \theta(d\gamma) = \delta \int_{t_1}^{t_2} \left(p_i \frac{dq_i}{dt} - H\right) dt = \delta \int_{t_1}^{t_2} L(q^i, \dot{q}^i, t) dt = 0.$$  \hfill (24)

IV. DISCUSSION AND CONCLUSIONS

We have found that the principle of stationary action is equivalent to assumptions (DR), (IND) and (KE) and that the variation of the action can be understood as the flow of evolutions through the surface delimited by the path and its variation. This allows us to fully understand both the physical and geometric significance of the principle in a much more precise way.

Note that the true physical content is in the displacement field $\dot{S}$ and the counting form $\omega$ whose properties descend from the assumptions. The action, the Lagrangian and the potential $\theta$ are not uniquely defined and are subject to a convenient, though still arbitrary, choice of gauge. The Lagrangian and the action, then, do not directly encode physical properties of the system. Saying “nature chooses the path that minimizes the action”, therefore, does not provide clear insight. On the other hand, the three assumptions give us a clear physical picture.

Another interesting element is that a version of the principle of stationary actions holds even if (KE) does not apply. In this case the “Lagrangian” $L(q, \dot{q}, p, t)$ depends on conjugate momentum as well. That is, the principle holds on all Hamiltonian systems, whether or not they admit a Lagrangian. Conversely, the cases where the Lagrangian does not admit a Hamiltonian are exactly the cases where the principle of stationary action fails to yield a unique solution, thus does not hold. These are the cases where the Lagrangian is not hyperregular. Therefore the principle of stationary action holds for all Lagrangian systems that admit a Hamiltonian, but also for all Hamiltonian systems whether or not they admit a Lagrangian. In this sense, the action principle is better understood as a feature of Hamiltonian mechanics, instead of Lagrangian mechanics [8, 10].

As a final comment, note that privileging the Hamiltonian picture is more in line with quantum mechanics; the idea of $\omega$ as a tool to count states is in line with statistical mechanics; expression 19 is remarkably similar to that of the relativistic four-momentum. This is one of the key insights we get from our Reverse Physics approach: there is a unity among the different physical theories that begs to be brought to light. We are convinced that a version of this geometric understanding must exist in the quantum world.

ACKNOWLEDGMENTS

This paper forms part of the ongoing Assumptions of Physics project [15], which aims to identify a handful of physical principles from which the basic laws can be rigorously derived. The authors would like to thank Alejandro Uribe and Daniel Burns for helpful discussions.
SUPPLEMENTARY INFORMATION

In this Supplementary Information we will review the arguments in more detail. Given that this article is still aimed at a physics audience, the focus will be conceptual rigor, rather than mathematical rigor. What is important is to establish a precise map between physical concepts and their mathematical representation, so that it is clear what physical requirement is captured by each mathematical definition.

We divide the material into three sections: the first will address the single d.o.f. case, the second will give a geometric interpretation of the tools of differential topology that is suitable for physicists, and lastly we address the multiple d.o.f. case.

One degree of freedom

For the case of one degree of freedom we are going to use vector calculus. This is possible since the problem can be studied with three variables (position, momentum and time)

\[ \xi^a = [q, p, t] \]  \hspace{1cm} (1)

which is the only setting where vector calculus truly works. In this sense, nature gave us a hand as three dimensional spaces are much more intuitive for us, and we can therefore develop the ideas in a familiar context before generalizing.

We break down the problem into sequential steps, which serves as a summary for the whole argument. We will have to:

(i) show that assuming determinism and reversibility (DR) is equivalent to assuming the existence of the displacement field \( \tilde{S} \) defined in 2 and its being divergence-free as in 4
(ii) show that the vector potential \( \tilde{\theta} \) for \( \tilde{S} \) defined in 5 can be expressed, without loss of generality, as 6
(iii) show that 4 and 6 recover Hamilton’s equations as in 7
(iv) show that the flow of \( \tilde{S} \) through a surface \( \Sigma \) delimited by a path \( \gamma \) and its variation \( \gamma' \) is the variation of the integral of the vector potential, as claimed in 8
(v) show that the integral of the vector potential is stationary for and only for the solutions of the equations of motion
(vi) show that the integral of the vector potential reduces to the action if and only if (KE) applies.

For (i), determinism and reversibility means that giving the state at some point in time means constraining the state at all times, that the system can evolve in only one way. In math terms, if \( X \) is the manifold of all possible states at all times, given \( P = (q_0, p_0, t_0) \) there will be only one possible evolution \( \gamma : \mathbb{R} \rightarrow X \) such that \( \gamma(t_0) = P \).

If \( X \) were a discrete space, the previous mathematical condition would be sufficient, but on a manifold it is not. Consider a damped harmonic oscillator: it satisfies the previous condition, yet the system will concentrate at the equilibrium. What needs to happen is that densities also need to be mapped exactly. That is, if a certain fraction of the system is in a given initial state, the same fraction of the system must be in the corresponding final state. Equivalently, if the system has a given probability of starting in a given initial state, it must have the same probability of ending in the corresponding final state. In mathematical terms, if \( \rho : X \rightarrow \mathbb{R} \) is a statistical distribution (or a probability density) defined on our extended
phase space and $\gamma$ is a possible evolution, we must have $\rho(\gamma(t_0)) = \rho(\gamma(t_1))$.

This requirement tells us two things. First, that $X$ must be a differentiable manifold. To define densities over a set of variables and map them backward and forward in time Jacobians must be defined, which means the derivative across state variables must be defined as well, in particular with respect to time, which plays the double role of a coordinate of $X$ and the evolution parameter. This means that the evolutions must be differentiable curves and therefore the displacement $\dot{S}$ must be defined, which leads to

$$\dot{S} = \left[ \frac{dq}{dt}, \frac{dp}{dt}, \frac{dt}{dt} \right] = S^a e_a = \frac{dS^a}{dt} e_a.$$  \hspace{1cm} (2)

In the literature of dynamical systems, the displacement vector field $\dot{S}$ is simply called the vector field of the dynamical system. The expression of the vector field is what determines the equations of motion in a dynamical system. The double role of time gives us

$$\dot{S} = \frac{dt}{dt} = 1. \hspace{1cm} (3)$$

Secondly, since the densities must be constant over evolutions, the flow of these densities through a closed surface in $X$ must be zero: each evolution will enter and exit the region delimited by the surface, giving no net contribution. Mathematically this requires the displacement to be divergence-free, which leads to

$$\nabla \cdot \dot{S} = \partial_a S^a = 0. \hspace{1cm} (4)$$

For (ii), since $\dot{S}$ is divergence-free, we can find a vector potential

$$\check{\theta} = \left[ \theta_q, \theta_p, \theta_t \right] = \theta_a e^a$$

$$\dot{S} = -\nabla \times \check{\theta} = -\epsilon^{abc} \partial_b \theta_c e_a.$$  \hspace{1cm} (5)

Note that the vector potential is not uniquely identified, since $\nabla \times (\check{\theta} + \nabla f) = \nabla \times \check{\theta}$ and therefore the displacement field $\dot{S}$ remains unchanged. This is commonly referred to as gauge freedom in physics, and it is analogous to what happens for a magnetic field or for an incompressible fluid.

We use this gauge freedom to set the momentum component to zero. In fact, we can always choose $f(q,p,t)$ such that $\partial_p f = -\theta_p$ and redefine $\check{\theta}$ to be $\check{\theta} + \nabla f$. Therefore we can set

$$\check{\theta} = \left[ \theta_q, 0, \theta_t \right] \hspace{1cm} (25)$$

without loss of generality.

We now use constraint $3$ on the time component. We have

$$S^t = \frac{dt}{dt} = -\left( \frac{\partial}{\partial q} \theta_p - \frac{\partial}{\partial p} \theta_q \right) \hspace{1cm} (26)$$

$$= -\left( \frac{\partial g}{\partial q} \right) \frac{\partial \theta_q}{\partial q} - \frac{\partial \theta_p}{\partial p} \hspace{1cm} (27)$$

$$= \frac{\partial \theta_q}{\partial p} = 1 \hspace{1cm} (28)$$

Integrating, we have $\theta_q = p + g(q,t)$ where $g(q,t)$ is an arbitrary function. This function can be set to zero without loss of generality as it can be removed with a gauge transformation where $f(q,t)$ does not depend on $p$, and therefore $\theta_p$ will remain unchanged. Therefore we have:

$$\check{\theta} = [p, 0, \theta_t] \hspace{1cm} (29)$$

Lastly, we rename the last component as $\theta_t = -H$ which leads to

$$\check{\theta} = [p, 0, -H] \hspace{1cm} (30)$$

$$= pe^t - H e^t.$$  \hspace{1cm} (6)

The potential $\check{\theta}$ is closely related to the canonical one-form of symplectic geometry and the contact form of contact geometry, but there are some differences. Both the canonical one-form and the contact form are a specific covector, defined with specific coordinates. For us, $\check{\theta}$ is a potential and therefore is not fully specified by the physics. Writing down a specific potential means choosing a gauge, therefore our potential corresponds to the canonical one-form or the contact form only if a specific gauge is fixed.

For (iii), we expand each component of $5$ using $6$. Again note the double role of time: $\frac{d\theta}{dt}$ is not the same as $\frac{d\theta}{\partial t}$. In the first case we are taking a total derivative along the evolution, and therefore the momentum can change. In the second case we are taking a partial derivative at constant $q$ and $p$ by definition, and therefore $\frac{d\theta}{\partial t} = 0$. We have

$$S^q = \frac{dq}{dt} = -\left( \frac{\partial}{\partial p} \theta_t - \frac{\partial \theta_q}{\partial t} \right)$$

$$= -\left( \frac{\partial g}{\partial q} \right) \frac{\partial \theta_q}{\partial q} - \frac{\partial \theta_p}{\partial p} \frac{\partial \theta_q}{\partial q}$$

$$= \frac{\partial H}{\partial p} \hspace{1cm} (30)$$

$^5$ We could be more precise and use $\tau$ for the evolution parameter, leaving $t$ just for the coordinate. The treatment would automatically acquire a relativistic flavor and would enlighten us to other interesting results, though it would distract from the main focus of this paper, the action principle.
Therefore we can write \( d\Sigma = d\hat{\gamma} \times d\hat{\lambda} \), where \( d\hat{\lambda} \) is the infinitesimal displacement that connects \( \gamma \) with \( \gamma' \). The integrand then becomes the triple product

\[
\hat{S} \cdot d\hat{\gamma} \times d\hat{\lambda}.
\]

We can understand it geometrically as the volume element given by the displacement \( \hat{S} \), the infinitesimal line element \( d\hat{\gamma} \) and the infinitesimal variation \( d\hat{\lambda} \) that moves \( \gamma \) to \( \gamma' \).

If \( \gamma \) is a solution of the equations of motion, \( \gamma \) is a field line of \( \hat{S} \): \( \gamma \) is tangent to \( \hat{S} \) at all points. We can write \( d\hat{\gamma} = \hat{S} dt \). The triple product, in this case, is zero regardless of what \( d\hat{\lambda} \) is. We can reason the other way: we look for a line \( \gamma \) such that the triple product is zero regardless of what \( d\hat{\lambda} \) is. The only way for that to happen is if \( \gamma \) is tangent to \( \hat{S} \) at each point, which means \( \gamma \) is a solution of the equations of motion. This means that the paths that make the line integral of the vector potential stationary are those and only those that solve the equations of motion. This recovers the principle of stationary action in Hamiltonian form

\[
0 = \delta \int_{\gamma} \hat{\theta} \cdot d\hat{\gamma} = - \iint_{\Sigma} \hat{S} \cdot d\Sigma.
\]

We want to stress that this variational principle applies to any divergence-free field. For example, for a magnetic field we could write:

\[
\iint_{\Sigma} \hat{B} \cdot d\Sigma = \delta \int_{\gamma} \hat{A} \cdot d\hat{\gamma}.
\]

We would find that the field lines of \( \hat{B} \) are the lines and only the lines for which the integral of the vector potential is stationary. Similarly, we would find that the streamlines of an incompressible fluid are the lines and only the lines for which the integral of the vector potential of the flow is stationary. While these properties may be of limited practical use, the geometric insight is the same.

For (vi), if the assumption (KE) of kinematic equivalence holds, we can from a trajectory in space \( q(t) \) reconstruct the state at any time. That is, if we are given \( q(t) \) we can reconstruct the full evolution \( \gamma(t) \) in the extended phase space.

We saw before that the evolution must be differentiable, which means \( q(t) \) must also be differentiable. Therefore the velocity \( \dot{q} = \frac{dq}{dt} \) is defined at every instant. Using assumption (DR), we know that each evolution is fully identified by position and momentum at a given time. Therefore the space of all possible evolutions is two-dimensional. Given that for each evolution \( \gamma(t) \) there is one and only one spatial trajectory \( q(t) \), the space of all spatial trajectories must also be two-dimensional and therefore fully identified by position and velocity at a given instant. Therefore there must be a map \( p(q, \dot{q}, t) \) that allows us to reconstruct the momentum at a given time by knowing position and velocity at that time.
We can use this map to express the Hamiltonian form of the principle of stationary action only in terms of position and velocity. We have

\[
\delta \int_{\gamma} \dot{\theta} \cdot d\gamma = \delta \int_{t_1}^{t_2} \dot{\theta} \cdot \frac{d\gamma}{dt} dt
= \delta \int_{t_1}^{t_2} \left( \theta_q \frac{dq}{dt} + \theta_p \frac{dp}{dt} + \theta_t \frac{dt}{dt} \right) dt
= \delta \int_{t_1}^{t_2} \left( p \frac{dq}{dt} + 0 \frac{dp}{dt} - H \frac{dt}{dt} \right) dt
= \delta \int_{t_1}^{t_2} \left( p \frac{dq}{dt} - H \right) dt
= \delta \int_{t_1}^{t_2} L(q, \dot{q}, t) dt = 0
\]  

(10)

We recognize the principle of stationary action in standard form.

Note that the argument can proceed in the opposite way. If we have a system for which the principle of stationary action yields a unique solution, then we can write \( p = \frac{\partial L}{\partial \dot{q}} = p(q, \dot{q}, t) \) which means we are able to reconstruct the momentum, and therefore the state, at each time. The system satisfies (KE). Moreover, we can find a unique Hamiltonian such that the solutions of the equations of motion are exactly those paths that satisfy the principle of stationary action. Therefore the system satisfies (DR) as well.

Before proceeding to the generalization, a few observations are in order.

- We stress that the vector potential \( \tilde{\theta} \) is not uniquely defined, and therefore neither are the action nor the Lagrangian. Giving precise physical characterizations of those two objects, then, is in our view a hopeless task.
- We found that the Hamiltonian \( H \) is the time component of the vector potential \( \tilde{\theta} \), not a scalar. This relativistic aspect emerges naturally out of the assumptions.
- One can, in principle, write Lagrangian systems for which the action principle does not yield a single solution (e.g. for \( L = 0 \) all paths are stationary). Consistently, assumption (DR) fails.

**Introduction to differential topology**

The generalization to multiple d.o.f will require the use of differential topology. These tools are not typically taught in a standard physics curriculum. If they are, their presentation is dense and abstract. Apart from the abstract definitions, the notation gets in the way of the physics. For example, writing \( \frac{\partial}{\partial \gamma} \) for the vector basis makes sense if one is studying the space of linear operators defined by derivations. When studying the space of velocities, however, writing \( \tilde{\phi} = v^i \frac{\partial}{\partial x^i} \) decreases understanding. Therefore we will use a notation and nomenclature that is closer to what physicists already use (e.g. Einstein notation) and need to picture. For example, we will use \( e_i \) for vector basis and \( e^i \) for covector (dual) basis. To those that believe we should just use the “correct” mathematical nomenclature, we retort that there is no such thing. Notations and names are chosen within a certain context to make the points of the discourse more clear.

What we will give here, then, is not a fully rigorous introduction to the subject, but rather a geometric understanding of what these tools describe and a physical motivation as to why they are useful. We hope this will make it easier to follow the next section and to build a geometric and physical intuition of the principle of stationary action, which is the focus of this article.

The core objective of using differential topology in physics is to characterize those quantities that are associated with a region of space and can be understood as sums of independent contributions. The mass \( m \) within a region \( V \) can be understood as the sum of the contributions of a mass density \( \rho \) within each infinitesimal volume \( dV \):

\[ m(V) = \iiint_V \rho dV. \]

The magnetic flux \( \Phi \) through a surface \( \Sigma \) can be understood as the sum of the contributions of the magnetic field \( \vec{B} \) over each infinitesimal \( d\Sigma \):

\[ \Phi(\Sigma) = \iint_{\Sigma} \vec{B} \cdot d\Sigma. \]

The work \( W \) over a path \( \gamma \) can be understood as the sum of the contributions of a force field \( \vec{f} \) over each infinitesimal line \( d\gamma \):

\[ W(\gamma) = \int_{\gamma} \vec{f} \cdot d\gamma. \]

Though the dimensionality changes, a pattern emerges: the functionals \( W, \Phi \) and \( m \) all take a region of space while \( \vec{f}, \vec{B} \) and \( \rho \) act on infinitesimal regions. However, the vector representation is in terms of different objects (i.e. vector, pseudo-vector and pseudo-scalar) and operations. Thus it fails to provide a nice generalization.

Suppose we wrote

\[
W(\gamma) = \int_{\gamma} f(d\gamma) \\
\Phi(\Sigma) = \iint_{\Sigma} B(d\Sigma) \\
m(V) = \iiint_V \rho(dV).
\]

Here we don’t have vectors, but functions of infinitesimal regions, which we call differential forms. The force, in this notation, is a one-form (or covector) as it takes one dimensional infinitesimal regions (i.e. vectors); the magnetic field is a two-form; the density is a three-form. We can also say that a scalar field, like the temperature, is a zero-form.
An infinitesimal region can also be understood as a parallelepiped, which is fully identified by its sides. Therefore a differential form can be understood as acting on a set of vectors, the sides of the parallelepiped, that matches the dimensionality of the form. A one-form will take one vector, a two-form two vectors and so on. All forms must be anti-symmetric because switching the order of the sides does not change the parallelepiped, just its orientation.

This picture is much easier to generalize to any dimensionality. If we note $S^k$ the space of all the $k$ dimensional subregions, a $k$-functional $F : S^k \to \mathbb{R}$ is a linear functional that takes a $k$-surface $\sigma$ and returns a number. This can be expressed as

$$F(\sigma) = \int_\sigma \omega(d\sigma),$$

the integral of a $k$-form $\omega : V^k \to \mathbb{R}$ that takes an infinitesimal $k$-surface $d\sigma$ (or equivalently a set of $k$ vectors) and returns a number.

This allows us to identify properties and theorems that apply to all functionals and forms regardless of their dimensionality. For example, if $\sigma_1$ and $\sigma_2$ are two disjoint regions, $F(\sigma_1 \cup \sigma_2) = F(\sigma_1) + F(\sigma_2)$, which also tells us $F(\emptyset) = 0$ for any functional. Additionally, $k$-forms are linear maps and can be written in terms of the tensor product of the co-basis $e^a$. We have:

$$f = f_a e^a$$
$$B = B_{abc} e^a \otimes e^b$$
$$\rho = \rho_{abc} e^a \otimes e^b \otimes e^c.$$

These tensors are all anti-symmetric as the forms are anti-symmetric.

We can see how this works for a force applied along a path. The infinitesimal displacement along $\gamma$ is given by

$$d\gamma = dx^a e_a.$$

Therefore

$$W(\gamma) = \int_\gamma f(d\gamma) = \int_\gamma f_a e^a dx^b e_b$$
$$= \int_\gamma f_a dx^b e^a(e_b) = \int_\gamma f_a dx^b \delta^a_b = \int_\gamma f_a dx^a.$$

Of all generalized properties, the most impressive one is certainly the generalized Stokes’ theorem. Compare the divergence (or Gauss’) theorem:

$$\mathcal{I} \mathcal{I} \mathcal{I} V (\nabla \cdot \vec{B}) dV = \mathcal{I} \mathcal{I} \mathcal{I} \Sigma_{\partial V} \vec{B} \cdot d\Sigma$$

with Stokes’ theorem

$$\mathcal{I} \Sigma (\nabla \times \vec{f}) \cdot d\Sigma = \mathcal{I} \gamma_{\partial \Sigma} \vec{f} \cdot d\gamma$$

and the divergence theorem

$$\mathcal{I}_\gamma (\nabla \varphi) \cdot d\gamma = [\varphi]_A^B.$$

In all three cases we have a region where we define the integral of some differential operator applied to a field. This is equated to the integral over the boundary of the field itself. But again, the vector calculus notation prevents us from generalizing as the differential operator is somewhat different in each of the three cases.

To understand conceptually that a generalization must exist, suppose we have a $k$-functional $F$ and its corresponding $k$-form $\omega$. By definition $F$ acts on $k$-surfaces. Now suppose we have a $k+1$-dimensional surface $\Sigma$. This is not a valid argument for $F$. However, its boundary $\partial \Sigma$ will be a $k$-dimensional surface, which is a valid argument for $F$. Therefore from any $F$ we can create a $k+1$-functional, which we can call the exterior functional of $F$, noted $\partial F$, such that $\partial F(\Sigma) = F(\partial \Sigma)$. If we look at all three cases above, that is exactly what is happening: the gradient, the curl and the divergence lead to a functional one dimension higher by applying the original operation on the boundary.

Now, since every functional has a corresponding form, $\partial F$ must have a corresponding $k+1$-form which must be fully identified by $\omega$. We note it as $\partial \wedge \omega$ and write

$$\partial F(\Sigma) = \int_\Sigma \partial \wedge \omega(d\Sigma) = \int_{\partial \Sigma} \omega(d\sigma) = F(\partial \Sigma).$$

This is the generalized Stokes’ theorem and $\partial \wedge \omega$ is the exterior derivative of $\omega$. We use $\partial \wedge$ as we leave $d$ reserved for the infinitesimal regions.

With a few steps, one can see that the divergence, curl and gradient correspond to the exterior derivative in three dimensions for a zero-form, one-form and two-form respectively, and that the three theorems of vector calculus are particular instances of this generalized one.

Moving on, the identities

$$\nabla \cdot \nabla \times \vec{f} = 0$$
$$\nabla \times \nabla \varphi = 0$$

both correspond to the single identity

$$\partial \wedge \partial \wedge \omega = 0,$$

which states that the exterior derivative applied twice is also zero. This identity can be understood in terms of the functionals. We have $\partial \partial F(\Sigma) = F(\partial \Sigma)$ and the boundary of a boundary $\partial \partial \Sigma = \emptyset$ is always the empty set. So, no matter what $F$ or $\Sigma$ is, we are evaluating a functional on an empty set, which is zero.

In the same way that a divergence-free field can be written as the curl of a vector potential, or a curl-free field can be written as the gradient of a scalar potential, a form whose exterior derivative is zero $\partial \wedge \omega = 0$ (a closed form), can be written as the exterior derivative $\omega = \partial \wedge \theta$ of a lower dimensional form, at least in some neighborhood. In terms of functionals, this corresponds to the case where $F(\sigma) = 0$ for all contractible surfaces, those surfaces that can be morphed to a point with a continuous transformation. If $F(\sigma) = 0$ over all surfaces, instead, then the corresponding form is exact.
The overall point here is that understanding differential topology in terms of functionals of regions gives us much better motivation and intuition for the math.

**Multiple degrees of freedom**

The case of multiple degrees of freedom proceeds in the same way as the single degree of freedom. The main difference is that we will work with a system described by $N$ degrees of freedom, which is therefore described by the $2N + 1$ variables

$$\xi^a = [q_i, p_i, t]. \quad (11)$$

We will have to:

(i) show that assuming determinism and reversibility (DR) and independence of degrees of freedom (IND) is equivalent to assuming the existence of the displacement field $\vec{S}$ defined by 12 and the closed counting form $\omega$ defined by 13 such that $\omega$ is killed only by the displacement field $\vec{S}$ as in 18 or by a field everywhere tangent to $\vec{S}$.

(ii) show that the potential $\theta$ for $\omega$ defined in 17 can be expressed, without loss of generality, as 19 and that $\omega$ can be expressed as 20.

(iii) show that 17, 18 and 19 recover Hamilton’s equations as in 21.

(iv) show that the integral of $\omega$ through a surface $\Sigma$ delimited by a path $\gamma$ and its variation $\gamma'$ is the variation of the integral of the vector potential, as claimed in 22.

(v) show that the integral of the vector potential is stationary for and only for the solutions of the equations of motion.

(vi) show that the integral of the vector potential reduces to the action if and only if (KE) applies.

For (i), the arguments presented for the single d.o.f. still apply: (DR) leads to the existence of a displacement field $\vec{S}$

$$\vec{S} = S^a e_a = \frac{d\xi^a}{dt} e_a = \frac{dq^i}{dt} e_{q^i} + \frac{dp_i}{dt} e_{p_i} + \frac{dt}{dt} e_t \quad (12)$$

that tells us how states move in the extended phase space. We also need to reformulate state/density conservation in time, which we cannot do simply with a surface integral of $\vec{S}$. To this end, we need a functional $F$ that takes a hyper-surface $\Sigma$ in the extended phase space and returns the count of states. We have

$$F(\Sigma) = \int_\Sigma \Omega(d\Sigma) \quad (35)$$

where $\Omega$ is the related form. Given the one-to-one correspondence between states and evolutions under (DR), $\Omega$ can be understood as quantifying either the flow of states through the surface or the states that are present on the surface.

As in the previous case, assumption (DR) tells us that the integral over a closed hyper-surface is zero, since as many states flow in as flow out. We have

$$\int_{\Sigma = \partial V} \Omega(d\Sigma) = 0 = \int_V \partial \wedge \Omega(dV) \quad (36)$$

which tells us that $\Omega$ is closed.

Given that $\Omega$ quantifies the flow of $\vec{S}$ through $\Sigma$, the flow for any surface that is tangent to $\vec{S}$ must vanish. If $d\gamma = \vec{S}\varphi dt$ for some scalar $\varphi$, we must have

$$\Omega(d\gamma, \cdot) = \varphi dt \Omega(\vec{S}, \cdot) = 0. \quad (37)$$

The notation is a bit improper as $\vec{S}$ is not technically an infinitesimal displacement, but a rate of displacement given a parameter. However, the slight abuse of notation makes the concept more clear: the direction of motion does not contribute to the flow through the surface $\Sigma$.

On the other hand, this must be the only direction that does not contribute to the flow. Since $\vec{S}$ cannot be zero because states must flow in time, for any other direction $\vec{T}$ that is not parallel to $\vec{S}$ there must be a surface tangent to $\vec{T}$ and not $\vec{S}$. The flow through that surface cannot be zero. Therefore the only way that the flow is zero for any surface tangent to a vector $\vec{T}$ is if $\vec{T} = \varphi \vec{S}$.

We now turn to assumption (IND). This requires the count of states to factorize: if we have a parallelepiped in phase space, the count of states (i.e. its volume) is the product of the count of the configurations over each independent d.o.f. (i.e. the area of the sides). Equivalently, (IND) must allow the case of a distribution over states with no correlations between the independent d.o.f., in which case the density must factorize. Since we must be able to quantify the configuration for each d.o.f., and since each d.o.f. is two dimensional, we must have a two-form $\omega$ such that:

$$\Omega = \omega^N. \quad (38)$$

Using assumptions (DR) and (IND) again we can see how $\omega$ must be closed

$$\partial \wedge \omega = 0. \quad (39)$$

If we imagine translating a surface along an independent variable, the number of configurations identified by the surface cannot change. This uses (IND). If we imagine translating in time, again the number of configurations cannot change because of (DR). Therefore, if we imagine a parallelepiped, opposite sides must identify the same number of configurations. The integral will take the sides with an opposite orientation, and therefore there will be zero net contribution. It is interesting to note that this condition requires both assumptions.

The form $\omega$ corresponds to the symplectic form of symplectic manifolds and of the symplectic bundles of contact manifolds with the following caveats. As we are in phase space extended by time, we have a contact manifold, and
ω mathematically corresponds to the exterior derivative of
the contact form, which is the primitive mathematical
object. Physically, this is backwards: ω is the actual prim-
itive physical object and the contact form θ is a potential,
which is unphysical as it is fixed by an arbitrary gauge.
Note how the mathematical definitions make it hard to
have a coherent physical interpretation. The primitive
objects for contact manifolds and symplectic manifolds
are different, θ in the first case and ω in the other, and
they are defined and studied as two separate mathemati-
cal spaces, while physically they are the same space with
or without time.

To summarize, assuming (DR) and (IND) gives us the
existence of the displacement vector field S and the closed
counting two-form ω such that S identifies the only vector
field that kills ω.

The converse is also true: if we start with the math-
ematical requirements, we can recover the physical as-
sumptions. If we are given a two-form ω on a 2N + 1
dimensional space, we can define the 2N form Ω = ω ∧ N.
This form will be non-zero because S is the only vector
field that kills the two-form ω. It will also be closed since
∂ ∧ ω = 0 and
\[ \partial ∧ Ω = N \partial ∧ ω ∧ ω^{N-1} = 0. \] (40)

So the flow of the well defined displacement field S, as
measured by Ω, is zero through closed surfaces. This im-
plies (DR).

To obtain (IND), if each pair of state variables q′, p_i
forms an independent degree of freedom, only the sur-
faces spanned by matching position and momentum will
define actual configurations, while all other combinations
will not define any. Therefore:
\[ \omega(e^{q_i}, e^{p_j}) = ω_{q_i,p_j} = δ^j_i = -ω_{p_j,q_i} \]
\[ \omega(e^{q_i}, e^{q_j}) = ω_{q_i,q_j} = 0 \]
\[ \omega(e^{p_i}, e^{p_j}) = ω_{p_i,p_j} = 0 \] (41)

If we start with a closed form ω that has a single de-
generate direction, by Darboux’s theorem we can find
coordinates for the non-degenerate part of ω that satisfy
the above conditions. This recovers (IND).

Note that we have broken down exactly how each
assumption contributes to the mathematical setting. If
(IND) fails, then Ω cannot be understood as ω^N, which
means we have fewer conditions on the motion: we only
know that the total flow is conserved, but we can-
not break it down along each d.o.f. On the other hand, if
(DR) fails we may not have a single displacement field S
and/or the form that counts states would not be closed.
Nonholonomic constraints will, in general, make one of
the assumptions fail, for example by introducing a de-
pendence between degrees of freedom. This type of exact
mapping between assumptions and their mathematical
expression is one of the key goals of Reverse Physics in
particular and our overall project Assumptions of Physics
more in general.

For (ii), we note that since the form is closed, we can
find, at least in a contractible region, a covector potential
θ such that ω = ∂ ∧ θ. Those familiar with symplectic
and contact geometry will know that using Darboux’s the-
orem, we can find a set of coordinates such that we can
express θ as
\[ \theta = p_i e^{q_i} - H e^t. \] (19)

Just stating this is, of course, not very insightful. To gain
more insight, we provide the sketch of a proof that, as we
have done for the single degree of freedom, uses gauge
transformations to fix the form of the potential.

As a starting point, we take conditions 41 that state
mathematically that q′ and p_i form independent degrees
of freedom. We re-express those conditions in terms of the
potential θ
\[ \omega(q_i', e^p_j) = (-∂ ∧ θ)_{q_i,p_j} = - (δ_{ij} p_i - δ_{ip} q_i) = δ^j_i \]
\[ \omega(e^{q_i}, e^{q_j}) = - (∂ᵢθⱼ - ∂ⱼθᵢ) = 0 \] (42)
\[ \omega(e^{p_i}, e^{p_j}) = - (∂ᵢθⱼ - ∂ⱼθᵢ) = 0 \]

We can use our gauge freedom to set θ_p_1 = 0, much
in the same way we did for the simpler case. We now have
∂ᵢθ_p = 0 and, by the first condition, ∂ₚθ_p = 1. Inte-
grating, we have θ_p = p₁ + q(g(q₁,p₂,p₃,...,t)) where
q is an arbitrary function which we can set to zero. Like in
the single degree of freedom, we can do this because g
does not depend on p₁, so it can be eliminated with a gauge
transformation that does not change θ_p_1. Therefore we
have:
\[ \theta = p₁ e^{q₁} + 0 e^{p₁} + q₁ e^{q₂} + ... + θₜ e^t. \] (43)

Note that the components for the first degree of free-
dom do not depend on the other degrees of freedom. That
is, for all i > 1, ∂ᵢθ_p = ∂ₚθ_p = ∂₀θ_p = ∂₁θ_p = 0,
But by using conditions 42, we find that the converse is
true as well: the components of all other degrees of free-
dom do not depend on the first. That is, for all i > 1,
∂ᵢθ_p = ∂ₚθ_p = ∂₀θ_p = ∂₁θ_p = 0.

We can then use, again, our gauge freedom with a func-
tion that does not depend on the first two variables to set
θ_p₂ = 0. And, with the same reasoning, we will be able
to set θ_q₂ = p₂. And then, again, find that the first two
degrees of freedom do not depend on the others, etc. At
the end, we will find 19.

The insight is that the independence of the degrees
of freedom leads to the independent expression of the
components of the covector potential. We stress again
that this is a choice, and that the specific expression of
the components is not physically constrained.
The components of $\partial \wedge \theta$ are

$$(\partial \wedge \theta)_{q'p_j} = \partial_q \theta_{p_j} - \partial_{p_j} \theta_q = \partial_q 0 - \partial_{p_j} p_i = -\delta^j_i$$

$$(\partial \wedge \theta)_{q'q'} = \partial_q \theta_{q'} - \partial_{q'} \theta_q = \partial_q p_j - \partial_{q'} p_i = 0$$

$$(\partial \wedge \theta)_{p_jp_j} = \partial_{p_j} \theta_{p_j} - \partial_{p_j} \theta_{p_j} = \partial_{p_j} 0 - \partial_{p_j} 0 = 0$$

$$(\partial \wedge \theta)_{q't} = \partial_q \theta_t - \partial_t \theta_q = \theta_q (-H) - \partial_t p_i = -\delta_q H$$

$$(\partial \wedge \theta)_{p_jt} = \partial_{p_j} \theta_t - \partial_t \theta_{p_j} = \theta_{p_j} (-H) - \partial t 0 = -\partial_{p_j} H$$

Therefore we recover

$$\omega_{ab} = (-\partial \wedge \theta)_{ab} = \begin{bmatrix} 0 & \delta^j_i & \partial_q H \\ -\delta^j_i & 0 & \partial_{p_j} H \end{bmatrix}. \quad (20)$$

For (iii) we calculate $\omega(\dot{S}, \cdot)$. We have

$$\omega(\dot{S}, \cdot) = S^a \omega_{ab} e^b = 0$$

$$= (S^q \dot{\omega}_{q'b} + S^p \omega_{pqb} + S^t \omega_{tqb}) e^b$$

$$= (S^q \dot{\omega}_{q'q'} + S^p \omega_{pq'b} + S^t \omega_{tq'b}) e^{q'} +$$

$$= (S^q \dot{\omega}_{q'p_j} + S^p \omega_{p_j q'} + S^t \omega_{t p_j q'}) e^{p_j} +$$

$$= (S^q \dot{\omega}_{q't} + S^p \omega_{p_j t} + S^t \omega_{t q'}) e^t$$

$$= (-S^p \delta^j_i - S^t \partial_q H) e^{q'} +$$

$$= (S^q \delta^j_i - S^t \partial_{p_j} H) e^{p_j} +$$

$$= (S^q \partial_q H + S^p \partial_{p_j} H) e^t$$

All components must be zero, therefore we have the following three equations:

$$S^p_j = -S^t \partial_q H$$

$$S^q = S^t \partial_{p_j} H$$

$$S^q \partial_{q'} H + S^p \partial_{p_j} H = S^q \partial_{p_j} H \partial_{q'} H - S^t \partial_q H \partial_{p_j} H = 0$$

(46)

Note that the last expression is not a new equation: it is identical to zero given the previous two equations. Since $S^t = 1$, we have

$$S^q = \partial_{p_j} H$$

$$S^p_j = -\partial_q H$$

(21)

which recovers Hamilton’s equations for multiple degrees of freedom.

The argument can be run in the other direction. If we have a system that obeys 21, we can define the two-form $\omega$ using 20. We can verify that $\omega$ is closed, it is killed by $\dot{S}$ and recover all the starting points. Therefore we find again complete equivalence with Hamiltonian mechanics and assumptions (DR) and (IND).

For those familiar with general relativity, recall that the metric tensor $g_{\alpha \beta}$ defines the geometry by defining lengths and angles. Here we find something similar in the sense that the two-form $\omega_{ab}$ is a rank two tensor that defines the geometry of the space. The difference is that $\omega_{ab}$ does not define lengths and angles, but only areas of two-dimensional surfaces, which corresponds to the number of configurations on each independent degree of freedom. Mathematically, this difference is between the symmetry of $g_{\alpha \beta} = g_{\beta \alpha}$ and the anti-symmetry of $\omega_{ab} = -\omega_{ba}$. In this context, what the equations of motion say is that states move in time at the same rate (i.e. $\frac{dt}{d\gamma} = 1$) with a deterministic and reversible flow (i.e. no states are created or destroyed on any independent degree of freedom). Nothing else.

Comparing 20 and 21 we see that the time components $\omega_{at}$ are actually the components of $\dot{S}$. In fact, the only non-trivial components of $\omega$ are components of $\dot{S}$. These two objects, then, are not only related: they express the same information.

For a single degree of freedom, recalling that $S^t = 1$, we can describe $\omega$ with the following nine components

$$\omega_{ab} = (-\partial \wedge \theta)_{ab} = \begin{bmatrix} 0 & S^t & -S^p \\ -S^t & 0 & S^q \\ S^p & -S^q & 0 \end{bmatrix}. \quad (47)$$

In three dimensions, using the Levi-Civita symbol $\epsilon^{abc}$ we can write

$$S^a = \epsilon^{abc} \omega_{bc}$$

(48)

which tells us that $\dot{S}$ is the pseudo-vector corresponding to the direction perpendicular to $\omega$. This is the same relationship that we would have, for example, for the magnetic field pseudo-vector and the spatial components of the electromagnetic field in relativity. If we express $\omega$ in terms of $\theta$, we have

$$S^a = -\epsilon^{abc} \partial_b \theta_c$$

(49)

which recovers 5: the displacement field is minus the curl of $\omega$. Thus we can see exactly how the general case reduces to the single d.o.f. case.

For (iv), we simply need to apply Stokes’ theorem. The situation is identical to the single degree of freedom, except that the flow now is not given by the surface integral of $\dot{S}$, but of the form $\omega$. The path $\gamma$ and its variation $\gamma'$ will form a closed curve where $\gamma'$ is taken with the opposite orientation. Because $\omega$ is closed, its surface integral depends only on its contour. Moreover, by Stokes’ theorem, the flow through the surface will be equal to the integral of the potential over the contour. Thus we have

$$-\int_{\Sigma} \omega(d\Sigma) = \int_{\Sigma} \partial \wedge \theta(d\Sigma)$$

$$= \oint_{\Sigma} \theta(d\gamma)$$

$$= \int_{\gamma'} \theta(d\gamma) - \int_{\gamma} \theta(d\gamma')$$

$$= \delta \int_{\gamma} \theta(d\gamma).$$

(22)

For (v), if the path $\gamma$ is a solution to the equations of motion, it is tangent to $\dot{S}$ at every point. Therefore we
can write $d\gamma = \vec{S} dt$. Each infinitesimal surface element $d\Sigma$ will be an infinitesimal parallelogram with $d\gamma$ as one side, and the other side $d\lambda$ connects the path to the variation. We have:

$$\delta \int_\gamma \theta (d\gamma) = - \int_\Sigma \omega (d\Sigma) = - \int_\Sigma \omega (d\gamma, d\lambda) = - \int_\Sigma \omega (\vec{S}, d\lambda) dt = 0. \quad (23)$$

The converse is also true: suppose that, for a fixed $\gamma$, the flow through $\Sigma$ is zero no matter the choice of $\gamma'$. Then at every point $\omega (d\gamma, \cdot) dt = 0$, which means $d\gamma$ kills $\omega$. But since $\vec{S}$ is the only direction that kills $\omega$ at every point, $d\gamma$ must be everywhere tangent to $\vec{S}$ and therefore $\gamma$ is a solution of the equations of motion.

For (vi), the argument is exactly the same as in the single d.o.f.. As the evolution is differentiable, the velocity $\dot{q}^i = \frac{dq^i}{dt}$ is well defined. Under (KE) there exists an invertible map $p_i = p_i (q^j, \dot{q}^j, t)$ that allows us to reconstruct the momentum at a given time by knowing position and velocity. We can therefore express the principle of stationary action only in terms of position and velocity. We have

$$\delta \int_\gamma \theta (d\gamma) = \delta \int_{t_1}^{t_2} \theta \left( \frac{d\gamma}{dt} \right) dt = \delta \int_{t_1}^{t_2} \left( \theta q^i \frac{dq^i}{dt} + \theta p_i \frac{dp_i}{dt} \right) dt = \delta \int_{t_1}^{t_2} \left( p_i \frac{dq^i}{dt} + H \frac{dt}{dt} \right) dt = \delta \int_{t_1}^{t_2} \left( p_i \frac{dq^i}{dt} - H \right) dt = \delta \int_{t_1}^{t_2} L(q^i, \dot{q}^i, t) dt = 0. \quad (24)$$

This recovers the principle of stationary action in standard form.

Again, the argument can proceed in the opposite direction. If we have a system for which the principle of stationary action yields a unique solution, we can write $p_i = \frac{\partial L}{\partial \dot{q}^i} = p_i (q^j, \dot{q}^j, t)$. This means that we can reconstruct the momentum from the kinematics, and therefore (KE) is satisfied. Moreover, we can define a Hamiltonian $H = p_i \dot{q}^i - L$ such that the paths that satisfy the equations of motion 21 are exactly those that satisfy the principle of stationary action. Therefore the system satisfies (DR) and (IND) as well.

The comments on the lack of strict physicality for the values of the Lagrangian and the action remain unchanged. They both depend on a choice of gauge. The comment on the centrality of the extended phase space to understand the action principle is only reinforced. Additionally, note that by changing $t$ to $ct$ expression 19 becomes

$$\theta = p_i e^{\theta^i} - \frac{H}{c} e^{ct}, \quad (50)$$

which bears striking resemblance to the relativistic four-momentum. While it is outside the scope of this article, the idea is that, in this setting, we already have some pre-relativistic features, even though we have not introduced a metric tensor.

We hope that this result shows that it is possible to bring physical ideas and mathematical concepts much closer together in a way that is more clear and satisfying.