

A Universal Variational Principle for Point Particle Classical Mechanics

W. A. Horowitz

*Department of Physics, University of Cape Town,
Private Bag X3, Rondebosch 7701, South Africa*

*Department of Physics, New Mexico State University, Las Cruces, New Mexico, 88003, USA and
Theoretical Sciences Visiting Program, Okinawa Institute of Science
and Technology Graduate University, Onna, 904-0495, Japan*

(Dated: May 4, 2026)

We derive from Newton's Second Law as an initial value problem initial—rather than boundary-value—action principles in configuration space and phase space for the classical motion of a point particle subject to any set of forces: conservative forces, holonomic and non-holonomic forces of constraint that do no virtual work, and general non-conservative forces. We provide explicit constructions of the actions given the arbitrary forces acting on the point particle. This work places action principles on equal footing with Newton's Second Law in classical mechanics.

I. INTRODUCTION

Newton's Second Law [1]

$$\sum_i \vec{F}_i \left(\left\{ \frac{d^I \vec{x}}{dt^I} \right\}_{I=0}^N; t \right) = m \frac{d^2 \vec{x}}{dt^2} \quad (1)$$

is the equation of motion for a classical point particle of mass m subject to arbitrary external forces that may depend on an arbitrary number of time derivatives of arbitrary order $I = 0, \dots, N$ of the path of the particle $\vec{x}(t)$ in some Cartesian inertial coordinate system. The Euler–Lagrange equations [2, 3],

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = \tilde{Q}_i^{\text{NC}} \left(\left\{ \frac{d^I \vec{q}}{dt^I} \right\}_{I=0}^N; t \right), \quad (2)$$

are equivalent to Newton's Second Law but have the advantage of working with a scalar quantity, the Lagrangian, and generalized coordinates. Generalized coordinates can be non-Euclidean, time dependent, etc., and have the significant advantage that one may eliminate redundant degrees of freedom associated with holonomic constraints—constraints that only depend on position—that do no virtual work. Conservative forces are incorporated into the Lagrangian while the applied forces are non-conservative forces, which may include forces of constraint. Recent work [4] has shown how to incorporate non-holonomic forces of constraint that may depend on any number of time derivatives of the path in the Euler-Lagrange framework. For non-holonomic constraints

$$h_k \left(\left\{ \frac{d^I \vec{q}}{dt^I} \right\}_{I=0}^M; t \right) = 0 \quad (3)$$

that do no virtual work, [4] tells us that

$$\tilde{Q}_i^{\text{NC}} \left(\left\{ \frac{d^I q}{dt^I} \right\}_{I=0}^N; t \right) = Q_i^{\text{NC}} \left(\left\{ \frac{d^I q}{dt^I} \right\}_{I=0}^N; t \right) + \lambda^k \frac{\partial h_k}{\partial \left(\frac{d^M q^i}{dt^M} \right)}, \quad (4)$$

where the λ^k are undetermined Lagrange multipliers.

Hamilton's Principle [5, 6] states that the physical path a particle takes—i.e. a path that satisfies the Euler–Lagrange Equations or, equivalently, Newton's Second Law—is the path that extremizes the action

$$S \equiv \int_{t_i}^{t_f} dt L[\vec{q}(t), \dot{\vec{q}}(t); t] \quad (5)$$

subject to variations in the path such that the initial and final positions of the particle, $\vec{q}(t_i)$ and $\vec{q}(t_f)$, are held fixed, i.e. subject to variations that satisfy

$$\delta \vec{q}(t_i) = \delta \vec{q}(t_f) = \vec{0}. \quad (6)$$

The variational principle approach has a number of advantages over approaches that start from the equations of motion, including making the role of symmetries transparent via Noether's Theorem; easily generalizing to continuous distributions of mass and to fields; and allowing for global rather than local solutions of classical paths. However, Hamilton's Principle suffers from a number of limitations. On a practical level, Eq. 5 is restricted to conservative forces and holonomic forces of constraint. As a result, while Newton's Second Law Eq. 1 or, equivalently, the Euler–Lagrange Equations Eq. 2 apply to any possible physical setup, Hamilton's Principle is a less fundamental, more limited theory. On a practical and conceptual level, Hamilton's Principle is a boundary value problem; the “initial” data provided are the position of the point particle at the initial time and the final time. On the other hand, the equations of motion from Newton's Second Law Eq. 1 or, equivalently, the Euler–Lagrange Equations Eq. 2 can be supplemented either by initial conditions, the position and velocity of the particle at the initial time, or by boundary conditions, the initial and final positions of the particle. So Hamilton's Principle is again more limited than the Newton's Second Law or, equivalently, the Euler–Lagrange approach. Note that this difference in boundary vs. initial conditions is, conceptually and practically, a major one. By the Picard–Lindelöf theorem, existence and uniqueness of solutions are guaranteed for initial value problems (that satisfy very mild conditions [7]), consistent with our experience of Nature, for which any experimental setup will yield a reproducible result. On the other hand, there can be zero, one, or many solutions to a boundary value problem [7], reflecting the difficulty in experimentally setting up a temporal boundary value problem. Conceptually, the theoretical prediction of the classical path given the temporal boundary data violates causality as generating the prediction for the full path requires the acausal knowledge of the future position of the particle. A related conceptual issue is the derivation of the Euler–Lagrange Equations from Hamilton's Principle: once derived from the boundary value problem of Hamilton's Principle, one generally solves the Euler–Lagrange Equations as an initial value problem rather than the boundary value problem in which the problem was initially posed. This changing of the problem from using boundary data to initial data is logically inconsistent. Finally, Hamilton's Principle has been limited to holonomic constraints, while the Euler–Lagrange approach could address holonomic constraints as well as non-holonomic constraints linear in velocities [8]. Recent work [4] has shown that it is not possible to incorporate general non-holonomic constraints into Hamilton's Principle. From the perspective of quantum field theory, our current best and most fundamental understanding of the basic building blocks of physics, and for which the action and the variational approach are fundamental, it is odd that in classical mechanics the variational approach is significantly limited compared to the equations of motion approach of Newton's Second Law or, equivalently, the Euler–Lagrange Equations.

A relatively recent Letter by Galley [9] postulated a variational action principle in classical mechanics that claimed to simultaneously address two of the three above issues, the formulation of a variational action principle in classical mechanics as an initial value problem and one that could handle non-conservative forces. A key ingredient in the postulated principle was a doubling of the degrees of freedom: instead of a single path, $\vec{q}(t)$, Galley considered two paths, $\vec{q}_1(t)$ and $\vec{q}_2(t)$. This doubling of paths was reminiscent of the Schwinger–Keldysh in-in approach in quantum mechanics, which was noted in the Letter, and also of Bateman's classical variational principle for dissipative systems that introduced a mirror system to absorb the energy lost to dissipation from the original system [10]. The action that Galley postulated was

$$S_G \equiv \int_{t_i}^{t_f} dt \left\{ L[\vec{q}_1(t), \dot{\vec{q}}_1(t); t] - L[\vec{q}_2(t), \dot{\vec{q}}_2(t); t] + K[\vec{q}_1(t), \dot{\vec{q}}_1(t), \vec{q}_2(t), \dot{\vec{q}}_2(t); t] \right\}, \quad (7)$$

where L is the usual Lagrangian for the system, which can incorporate conservative forces and depends on only the path 1 or the path 2, and a new term, K , that depends on both paths 1 and 2 and that can incorporate non-conservative forces.

Galley claimed that extremization of Eq. 7 under variations in both paths, subject to the four conditions

$$\delta \vec{q}_1(t_i) = \vec{0}, \quad (8a)$$

$$\delta \vec{q}_2(t_i) = \vec{0}, \quad (8b)$$

$$\delta \vec{q}_1(t_f) = \delta \vec{q}_2(t_f), \quad (8c)$$

$$\delta \dot{\vec{q}}_1(t_f) = \delta \dot{\vec{q}}_2(t_f), \quad (8d)$$

yields the classical equations of motion for $\vec{q}_+(t) \equiv \frac{1}{2}(\vec{q}_1(t) + \vec{q}_2(t))$ for a system under conservative and non-conservative forces.

Galley's postulated variational principle with doubled degrees of freedom has found uses across a wide array of physics. The Galley non-conservative action allows one to derive classical compact-binary dynamics from worldline quantum field theory [11], to compute dissipative effects and to organize the gravitational radiation reaction order-by-order for compact binary-inspirals [12–15], and is related to the classical limit of open quantum systems [16]. The doubled degrees of freedom approach provides an alternative treatment of dissipation compared to using fractional

derivatives [17, 18]. Action principles are particularly valuable when working with fields, and doubled degrees of freedom allow one access to effective theories of dissipative continuum systems [19–21].

However, the Galley variational principle Eq. 7 and 8 also suffers from a number of issues: the “physical limit” $\vec{q}_-(t) \equiv \vec{q}_1(t) - \vec{q}_2(t) \rightarrow \vec{0}$ is taken by hand; no general construction of K for any general non-conservative force was given; no rigorous derivation of Eq. 7 or Eq. 8 from first principles of either classical or quantum mechanics was given, where, in particular, the condition Eq. 8d cannot come from a usual path integral formulation that only involves integrations over paths (one would require a configuration space path integral over both paths *and* velocities); and a numerical implementation of Eq. 7 and 8 yielded paths with unphysical jumps at the endpoints [22], often a sign in numerical optimization of an overconstrained system.

In order to address the above issues with [9], we recently carefully derived the $\hbar \rightarrow 0$ limit of the expectation value of operators (such as position and momentum) for a point particle of mass m subject to conservative, position-dependent forces in non-relativistic quantum mechanics [23]. We found that the dominant contribution to the path integral extremizes what we called the Hamilton Revised action,

$$S_{HR} \equiv \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ L[\vec{q}_1(t), \dot{\vec{q}}_1(t)] - L[\vec{q}_2(t), \dot{\vec{q}}_2(t)] \right\} \quad (9)$$

subject to variations in the paths with

$$\frac{1}{2}(\vec{q}_1(t_i) + \vec{q}_2(t_i)) = \vec{q}_0, \quad (10a)$$

$$\vec{q}_1(t_f) = \vec{q}_2(t_f) \quad (10b)$$

fixed. Notice, crucially, the extra $\vec{p}_0 \cdot \vec{q}_-(t_i)$ term in Eq. 9 compared to Galley’s Eq. 7, also missing from Bateman’s treatment, where

$$\vec{q}_+(t) \equiv \frac{1}{2}(\vec{q}_1(t) + \vec{q}_2(t)), \quad (11)$$

$$\vec{q}_-(t) \equiv \vec{q}_1(t) - \vec{q}_2(t). \quad (12)$$

This extra $\vec{p}_0 \cdot \vec{q}_-(t_i)$ term comes explicitly from the quantum initial wavepacket, whose smearing in both position and momentum space is required by the Heisenberg Uncertainty Principle, and is necessary for setting the initial velocity condition. Unlike Eq. 8, the conditions for the variation of the path, by construction, are consistent with quantum mechanics. And an implementation of Eq. 9 and 10 numerically yielded perfectly smooth trajectories, consistent with Newton’s Second Law as an initial value problem. Interestingly, extremization of Eq. 9 subject to Eq. 10 yields “initial conditions” for $\vec{q}_-$ at the final time of $\vec{q}_-(t_f) = \dot{\vec{q}}_-(t_f) = \vec{0}$ and equations of motion that propagate those initial conditions backwards in time uniquely setting $\vec{q}_-(t) = \vec{0}$ for all time; i.e. the “physical limit” $\vec{q}_-(t) = \vec{0}$ is set by the dynamics of Eq. 9 and 10 rather than having to be set by hand.

In a recent work [24], inspired by [4, 9], we postulated an action

$$S \equiv \int_{t_i}^{t_f} dt \left\{ L[\vec{q}_1(t), \dot{\vec{q}}_1(t); t] - L[\vec{q}_2(t), \dot{\vec{q}}_2(t); t] + \lambda_-^k(t) g_k[\vec{q}_+(t), \dot{\vec{q}}_+(t); t] - \lambda_+^k(t) q_-^i(t) \frac{\partial g_k}{\partial \left(\frac{d\vec{q}_+^i}{dt} \right)} \right\}, \quad (13)$$

with doubled degrees of freedom whose extremization we thought reproduced the Lagrange–d’Alembert equations of motion for conservative systems subject to general, velocity-dependent constraints

$$g_k[\vec{q}_+(t), \dot{\vec{q}}_+(t); t] = 0 \quad (14)$$

that perform no virtual work, Eq. 4 with $Q^{\text{NC}} = 0$ and $M = 1$. The work predated [23], so the $\vec{p}_0 \cdot \vec{q}_-(t_i)$ term discovered in [23] is missing.

Two questions naturally arose out of the work of [23, 24]: 1) could Eq. 9 and 10 be derived more simply from first principles in classical mechanics, rather than from the much more complicated $\hbar \rightarrow 0$ limit of Schwinger–Keldysh quantum mechanics, and 2) could Eq. 9 and 10 be generalized to incorporate any set of applied and constraint forces? The purpose of this work is to show that the answer to both of these questions is yes. In the process, we will give an explicit construction of actions, in configuration space and phase space, whose extremization yields the classical equations of motion as an initial value problem for any set of applied forces and forces of constraint.

This manuscript is organized as follows. In Sec. II, we derive directly from Newton’s Second Law as an initial value problem a configuration space variational principle valid for a classical point particle of mass m subject to any conservative force, any holonomic force of constraint that does no virtual work, and any non-conservative force

$\vec{Q}^{\text{NC}}[\vec{q}_+(t), \dot{\vec{q}}_+(t); t]$ that depends on at most the first time derivative of the path. We find that the action Eq. 9 is modified to

$$S \equiv \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ L[\vec{q}_1(t), \dot{\vec{q}}_1(t); t] - L[\vec{q}_2(t), \dot{\vec{q}}_2(t); t] + \vec{q}_-(t) \cdot \vec{Q}^{\text{NC}}[\vec{q}_+(t), \dot{\vec{q}}_+(t); t] \right\}. \quad (15)$$

A few additional lines of work show that extremization of Eq. 15 by varying paths subject to Eq. 10 is equivalent to the extremization of

$$S = \int_{t_i}^{t_f} dt q_-^i(t) \left[-\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_+^i} + \frac{\partial L}{\partial q_+^i} + Q_i^{\text{NC}}[\vec{q}_+(t), \dot{\vec{q}}_+(t); t] + \left(p_{0,i} - \frac{\partial L}{\partial \dot{q}_+^i} \right) \delta(t - t_i) \right] \quad (16)$$

by varying paths $\vec{q}_\pm(t)$ subject to

$$\begin{aligned} \vec{q}_+(t_i) &= \vec{q}_0, \\ \vec{q}_-(t_f) &= \vec{0} \end{aligned} \quad (17)$$

fixed. In Sec. III, we confirm explicitly that variation of Eq. 15 subject to Eq. 10 and Eq. 16 subject to Eq. 17 give paths that satisfy Newton's Second Law as an initial value problem. In Sec. IV, we similarly derive from Newton's Second Law a phase space variational principle valid for the same types of forces as in Sec. II. We find the action is given by

$$S = \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt' \left\{ \vec{p}_1 \cdot \dot{\vec{q}}_1 - H[\vec{p}_1, \vec{q}_1; t'] - (\vec{p}_2 \cdot \dot{\vec{q}}_2 - H[\vec{p}_2, \vec{q}_2; t']) + \vec{q}_- \cdot \vec{Q}^{\text{NC}}[\vec{q}_+, \vec{p}_+; t'] \right\}, \quad (18)$$

where

$$H[\vec{p}_a, \vec{q}_a; t] \equiv p_{a,i} \dot{q}_a^i - L[\vec{q}_a, \dot{\vec{q}}_a; t], \quad a = 1, 2, \quad (19)$$

$$p_{a,i} \equiv \frac{\partial L}{\partial \dot{q}_a^i}. \quad (20)$$

The varied paths must satisfy Eq. 17; there are no restrictions on the variations of the momenta. An equivalent action to extremize is

$$S = \int_{t_i}^{t_f} dt \left\{ p_{-,i} \left[\dot{q}_+^i - \frac{\partial H}{\partial p_{+,i}} \right] - q_-^i \left[\dot{p}_{+,i} + \frac{\partial H}{\partial q_+^i} - Q_i^{\text{NC}}[\vec{q}_+, \vec{p}_+; t] - (p_{0,i} - p_{+,i}) \delta(t - t_i) \right] \right\} \quad (21)$$

subject to paths that satisfy Eq. 17; there are again no restrictions on the variations of the momenta.

In Sec. V we generalize to any non-conservative forces that may depend on an arbitrary number of arbitrary orders of time derivatives of the path of the classical particle. In order to inject the initial conditions of the higher derivatives that the force depends on, we must introduce new variables

$$\vec{q}_+^{(I)} \equiv \frac{d^I \vec{q}_+}{dt^I}, \quad I = 2, \dots, N, \quad (22)$$

through the introduction of Lagrange multipliers $\vec{\lambda}_-^{(I)}$ and the additional term

$$\sum_{I=2}^N \vec{\lambda}_-^{(I)} \cdot \left(\frac{d^I \vec{q}_+}{dt^I} - \vec{q}_+^{(I)} \right) \quad (23)$$

in the various actions Eq. 15, 16, 18 and 21. In Sec. VI we show how to incorporate constraints

$$h_k \left(\left\{ \frac{d^I \vec{q}}{dt^I} \right\}_{I=0}^M; t \right) = 0$$

by including the term

$$\int_{t_i}^{t_f} dt \left\{ q_-^i \lambda_+^k(t) \frac{\partial h_k}{\partial \left(\frac{d^M q_+^i}{dt^M} \right)} + \lambda_-^k(t) h_k \left[\left\{ \frac{d^I \vec{q}_+}{dt^I} \right\}_{I=0}^M; t \right] \right\} \quad (24)$$

in the action. In Sec. VII we conclude and provide an outlook.

II. DERIVATION OF THE ACTION

We start with

$$\begin{aligned}\vec{F} &= m \ddot{\vec{x}}, \\ \vec{x}(t_i) &= \vec{x}_0, \\ \dot{\vec{x}}(t_i) &= \vec{v}_0.\end{aligned}\tag{25}$$

The force can, in principle, depend on the path $\vec{x}(t)$, any time derivatives of the path, and on time:

$$\vec{F} = \vec{F}\left(\left\{\frac{d^I \vec{x}}{dt^I}\right\}_{I=0}^N, t\right).\tag{26}$$

If the force depends at most on the path and its first derivative, then, by the Picard–Lindelöf theorem, a solution to Eq. 25 exists and is unique. If the force depends on higher derivatives of the path, then additional initial conditions must be supplied in order for a solution to exist (and be unique). For simplicity, we will assume in this section that the force depends at most on the velocity of the path.

We may always decompose the force into contributions that are conservative,

$$\vec{F}^C \equiv -\vec{\nabla} V,\tag{27}$$

where we assume that the potential depends on position alone, $V = V(\vec{x})$, and forces that are non-conservative, $\vec{F}^{\text{NC}}$. We may further decompose the non-conservative forces into forces of constraint that perform no virtual work, $\vec{F}^{\text{cons}}$, $\vec{F}^{\text{cons}} \cdot \delta \vec{x} = 0$, where the variations $\delta \vec{x}(t)$ are consistent with the constraints, and all other forces, $\vec{F}^{\text{NC}}$. We may thus write

$$\vec{F} = \vec{F}^C + \vec{F}^{\text{NC}} + \vec{F}^{\text{cons}}.\tag{28}$$

In Eq. 28, we can consider multiple conservative forces, non-conservative forces, and forces of constraint, indicated by a sum over an index for each type of force, but we will suppress this generalization to avoid unnecessary notational clutter.

Define generalized coordinates

$$\vec{q} \equiv \vec{q}(\vec{x})\tag{29}$$

and generalized forces

$$Q_i \equiv F_j \frac{\partial x^j}{\partial q^i}.\tag{30}$$

Then the conservative generalized forces are given by

$$Q_i^C = -\frac{\partial V}{\partial q^i}.\tag{31}$$

In these generalized coordinates, some of the generalized forces may be forces of constraint that do no virtual work,

$$\vec{Q}^{\text{cons}} \cdot \delta \vec{q} = 0.\tag{32}$$

In the generalized coordinates, the forces of constraint that do no virtual work may come from (some of) the forces of constraint $\vec{F}^{\text{cons}}$ and/or the non-conservative forces $\vec{F}^{\text{NC}}$ in the original coordinate system, depending on the particulars of the generalized coordinates. Let us therefore decompose the non-conservative generalized forces into forces of constraint that do no virtual work and all other non-conservative generalized forces,

$$\tilde{Q}_i^{\text{NC}} = \tilde{F}_j^{\text{NC}} \frac{\partial x^j}{\partial q^i} = Q_i^{\text{NC}} + Q_i^{\text{cons}}.\tag{33}$$

Define

$$g_{ij} \equiv \delta_{k\ell} \frac{\partial x^k}{\partial q^i} \frac{\partial x^\ell}{\partial q^j}.\tag{34}$$

g_{ij} depends only on the generalized coordinates, $g_{ij} = g_{ij}(\vec{q})$, and can be thought of as the pullback of the metric on the original Cartesian coordinates onto the manifold of the generalized coordinates. Then the kinetic energy in the generalized coordinates is given by

$$T = \frac{1}{2} m \delta_{ij} \dot{x}^i \dot{x}^j = \frac{1}{2} m g_{ij} \dot{q}^i \dot{q}^j. \quad (35)$$

Define the Lagrangian in the generalized coordinates as

$$L \equiv T - V = \frac{1}{2} m g_{ij}(\vec{q}) \dot{q}^i \dot{q}^j - V(\vec{q}). \quad (36)$$

We restrict our analysis to generalized coordinates for which $\partial^2 L / \partial \dot{q}^i \partial \dot{q}^j$ is non-degenerate; consideration of degenerate Lagrangians is much more involved [25, 26]. Define the generalized momentum as

$$p_i \equiv \frac{\partial L}{\partial \dot{q}^i} = m g_{ij} \dot{q}^j. \quad (37)$$

In our generalized coordinates, Eq. 25 can be written as (see, e.g., [8])

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} &= Q_i^{\text{NC}}, \\ \vec{q}(t_i) &= \vec{q}_0, \\ \dot{\vec{q}}(t_i) &= \dot{\vec{q}}_0, \end{aligned} \quad (38)$$

where

$$\begin{aligned} \vec{q}_0 &\equiv \vec{q}(\vec{x}_0), \\ \dot{q}_0^i &\equiv \left. \frac{\partial q^i}{\partial x^j} \right|_{\vec{x}_0} \dot{x}_0^j. \end{aligned} \quad (39)$$

Note that in Eq. 38, if forces of constraint that do no virtual work $\vec{Q}^{\text{cons}}$ are present, then each of these forces must be used to eliminate a degree of freedom from the generalized coordinates, as otherwise Eq. 38 is unaware of the presence of the constraint.

Let us consider a slightly different problem. Consider now two paths, $\vec{x}_1(t)$ and $\vec{x}_2(t)$, subject to the same forces as in the original problem

$$\vec{F}_1 = m \ddot{\vec{x}}_1 \quad \text{and} \quad \vec{F}_2 = m \ddot{\vec{x}}_2, \quad (40)$$

where $\vec{F}_i = \vec{F}(\vec{x}_i, \dot{\vec{x}}_i; t)$. In order to uniquely specify the two solutions $\vec{x}_a(t)$, we need to provide four conditions. We will see that the correct four initial conditions to impose in order to reproduce the classical motion associated with Eq. 25 are

$$\begin{aligned} \frac{1}{2} [\vec{x}_1(t_i) + \vec{x}_2(t_i)] &= \vec{x}_0, \\ \frac{1}{2} [\dot{\vec{x}}_1(t_i) + \dot{\vec{x}}_2(t_i)] &= \dot{\vec{x}}_0, \\ \vec{x}_1(t_f) - \vec{x}_2(t_f) &= \vec{0}, \\ \dot{\vec{x}}_1(t_f) - \dot{\vec{x}}_2(t_f) &= \vec{0}. \end{aligned} \quad (41)$$

We call Eq. 41 initial conditions, even though two of the conditions are at the final time, for reasons that will become clear in a moment.

In the generalized coordinates, Eq. 40 becomes (see, e.g., [8])

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_a^i} - \frac{\partial L}{\partial q_a^i} = Q_{a,i}^{\text{NC}}, \quad a = 1, 2, \quad (42)$$

where

$$\frac{\partial L}{\partial q_a^i} \equiv \left. \frac{\partial L[q^i, \dot{q}^i; t]}{\partial q^i} \right|_{\substack{q^i = q_a^i \\ \dot{q}^i = \dot{q}_a^i}}, \quad \frac{\partial L}{\partial \dot{q}_a^i} \equiv \left. \frac{\partial L[q^i, \dot{q}^i; t]}{\partial \dot{q}^i} \right|_{\substack{q^i = q_a^i \\ \dot{q}^i = \dot{q}_a^i}}, \quad (43)$$

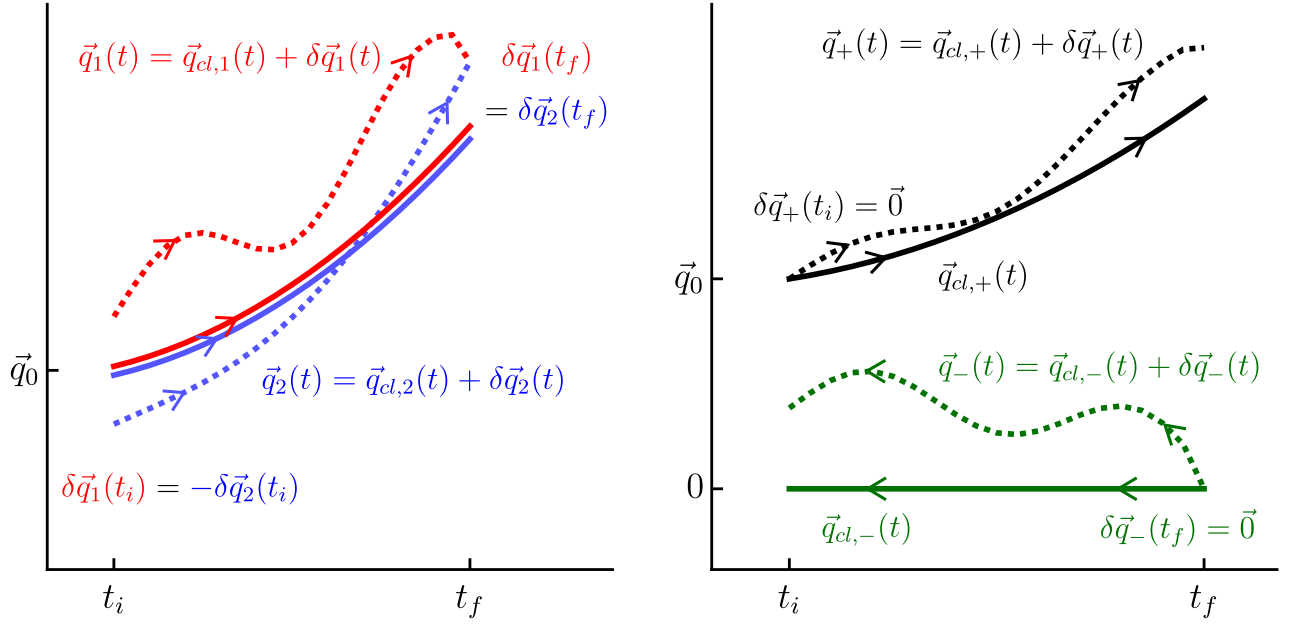


FIG. 1. (Left) The doubling of paths $\vec{q}_1(t)$ (red) and $\vec{q}_2(t)$ (blue). The classical paths are solid while the full paths including fluctuations are dashed. The path variations at t_i are equal and opposite while the variations at t_f are equal. Arrows indicate the forwards in time evolution of the paths. The classical paths are artificially slightly separated vertically for visualization purposes. (Right) The two paths $\vec{q}_+(t)$ (black) and $\vec{q}_-(t)$ (green) in the Keldysh basis. The classical paths are represented by the solid lines while the classical paths plus variations are represented by the dashed lines. The t_i position of the plus path is fixed while the t_f position of the minus path is fixed. The plus path evolves forward in time while the minus path evolves backwards in time.

and using Eq. 37, the initial conditions Eq. 41 are

$$\frac{1}{2} [\vec{q}_1(t_i) + \vec{q}_2(t_i)] = \vec{q}_0, \quad (44a)$$

$$\frac{1}{2} \left[\left. \frac{\partial L}{\partial \dot{\vec{q}}_1^i} \right|_{t_i} + \left. \frac{\partial L}{\partial \dot{\vec{q}}_2^i} \right|_{t_i} \right] = p_{0,i}, \quad (44b)$$

$$\vec{q}_1(t_f) - \vec{q}_2(t_f) = \vec{0}, \quad (44c)$$

$$\left. \frac{\partial L}{\partial \dot{\vec{q}}_1^i} \right|_{t_f} - \left. \frac{\partial L}{\partial \dot{\vec{q}}_2^i} \right|_{t_f} = 0, \quad (44d)$$

where

$$\begin{aligned} \vec{q}_0 &\equiv \vec{q}(\vec{x}_0), \\ p_{0,i} &\equiv m g_{ij}(\vec{q}_0) \dot{q}^j \Big|_{t_i} = m g_{ij}(\vec{q}_0) \frac{\partial q^j}{\partial x^k} \Big|_{\vec{x}_0} \dot{x}_0^k. \end{aligned} \quad (45)$$

Since we are working in configuration space—we will consider phase space later— $p_{0,i}$ should be considered as shorthand for setting the initial generalized velocity $\dot{q}^i(t_i)$ rather than for setting the initial momentum. When we transition to phase space, $p_{0,i}$ should then be considered as the initial momentum of the point particle.

We consider in the usual way paths that are classical and that include variations,

$$\vec{q}_a(t) = \vec{q}_{a,cl}(t) + \delta \vec{q}_a(t), \quad a = 1, 2. \quad (46)$$

From Eq. 44a and Eq. 44c, the variations at the initial and final times are not independent,

$$\begin{aligned} \delta \vec{q}_1(t_i) &= -\delta \vec{q}_2(t_i), \\ \delta \vec{q}_1(t_f) &= \delta \vec{q}_2(t_f); \end{aligned} \quad (47)$$

otherwise the variations are independent and arbitrary. See Fig. 1 (left).

Define δS as

$$\begin{aligned} \delta S \equiv & 2 \left[p_{0,i} - \frac{1}{2} \left(\left. \frac{\partial L}{\partial \dot{q}_1^i} \right|_{t_i} + \left. \frac{\partial L}{\partial \dot{q}_2^i} \right|_{t_i} \right) \right] \delta q_1^i(t_i) + \left[\left. \frac{\partial L}{\partial \dot{q}_1^i} \right|_{t_f} - \left. \frac{\partial L}{\partial \dot{q}_2^i} \right|_{t_f} \right] \delta q_1^i(t_f) \\ & + \int_{t_i}^{t_f} dt \left\{ \left(\frac{\partial L}{\partial q_1^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1^i} + Q_{1,i}^{\text{NC}} \right) \delta q_1^i - \left(\frac{\partial L}{\partial q_2^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_2^i} + Q_{2,i}^{\text{NC}} \right) \delta q_2^i \right\}. \end{aligned} \quad (48)$$

From the Euler-Lagrange equations Eq. 42 and the initial conditions Eq. 44b and 44d, each term multiplying a variation in a path is individually zero in Eq. 48. Hence we have that $\delta S = 0$ identically. From Eq. 47, we may write Eq. 48 as

$$\begin{aligned} 0 = \delta S = & \left[p_{0,i} - \left. \frac{\partial L}{\partial \dot{q}_1^i} \right|_{t_i} \right] \delta q_1^i(t_i) - \left[p_{0,i} - \left. \frac{\partial L}{\partial \dot{q}_2^i} \right|_{t_i} \right] \delta q_2^i(t_i) + \left[\left. \frac{\partial L}{\partial \dot{q}_1^i} \right|_{t_f} \right] \delta q_1^i(t_f) - \left[\left. \frac{\partial L}{\partial \dot{q}_2^i} \right|_{t_f} \right] \delta q_2^i(t_f) \\ & + \int_{t_i}^{t_f} dt \left\{ \left(\frac{\partial L}{\partial q_1^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_1^i} \right) \delta q_1^i - \left(\frac{\partial L}{\partial q_2^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_2^i} \right) \delta q_2^i + Q_{1,i}^{\text{NC}} \delta q_1^i - Q_{2,i}^{\text{NC}} \delta q_2^i \right\}. \end{aligned} \quad (49)$$

Define the following change of coordinates, often referred to as Keldysh coordinates,

$$\begin{aligned} \vec{q}_+(t) &\equiv \frac{1}{2} (\vec{q}_1(t) + \vec{q}_2(t)), & \vec{q}_1(t) &= \vec{q}_+(t) + \frac{1}{2} \vec{q}_-(t), \\ \vec{q}_-(t) &\equiv \vec{q}_1(t) - \vec{q}_2(t) & \vec{q}_2(t) &= \vec{q}_+(t) - \frac{1}{2} \vec{q}_-(t). \end{aligned} \quad (50)$$

Note that the change from the 1/2 paths to the $+/-$ paths is not entirely trivial. For example,

$$\begin{aligned} \frac{1}{2} \left(\frac{\partial L}{\partial \dot{q}_1^i} + \frac{\partial L}{\partial \dot{q}_2^i} \right) &= \frac{1}{2} \left(g_{ij}(\vec{q}_1) \dot{q}_1^j + g_{ij}(\vec{q}_2) \dot{q}_2^j \right) \\ &= \frac{1}{2} \left[g_{ij} \left(\vec{q}_+ + \frac{1}{2} \vec{q}_- \right) \left(\dot{q}_+^j + \frac{1}{2} \dot{q}_-^j \right) + g_{ij} \left(\vec{q}_+ - \frac{1}{2} \vec{q}_- \right) \left(\dot{q}_+^j - \frac{1}{2} \dot{q}_-^j \right) \right] \\ &= g_{ij}(\vec{q}_+) \dot{q}_+^j + \frac{1}{4} \frac{\partial g_{ij}}{\partial q_+^k} q_-^k \dot{q}_-^j + \dots \\ &= \frac{\partial L}{\partial \dot{q}_+^i} + \mathcal{O}(\vec{q}_-). \end{aligned} \quad (51)$$

We will find that the initial conditions Eq. 44 and the equations of motion Eq. 42 will yield $\vec{q}_-(t) = \vec{0}$ for all time; thus one will ultimately have that

$$\frac{1}{2} \left(\frac{\partial L}{\partial \dot{q}_1^i} + \frac{\partial L}{\partial \dot{q}_2^i} \right) = \frac{\partial L}{\partial \dot{q}_+^i}, \quad (52)$$

but that the equality in Eq. 52 was not automatic.

We will further find that $\vec{q}_+(t)$ will, once $\vec{q}_-(t)$ is set to zero, solve the classical equations of motion for the particle in generalized coordinates subject to initial conditions on the position and velocity in the generalized coordinates, Eq. 38, which is equivalent to finding a solution to Newton's Second Law subject to initial conditions on the path and velocity in the original coordinates, Eq. 25.

The initial conditions in the 1/2 coordinates, Eq. 44, thus become in the $+/-$ paths

$$\vec{q}_+(t_i) = \vec{q}_0, \quad (53a)$$

$$\left. \frac{\partial L}{\partial \dot{q}_+^i} \right|_{t_i} = p_{0,i}, \quad (53b)$$

$$\vec{q}_-(t_f) = \vec{0}, \quad (53c)$$

$$\left. \frac{\partial L}{\partial \dot{q}_-^i} \right|_{t_f} = 0, \quad (53d)$$

where, in Eq. 53b and Eq. 53d, we have already assumed that $\vec{q}_-(t) = \vec{0} \forall t$.

The $+/-$ coordinates, too, have classical paths and variations around the classical paths,

$$\vec{q}_{\pm}(t) = \vec{q}_{\pm,\text{cl}}(t) + \delta\vec{q}_{\pm}(t). \quad (54)$$

Similar to the $1/2$ coordinates, the variations in the $+/-$ coordinates are restricted by Eq. 53a and Eq. 53c,

$$\begin{aligned} \delta q_+^i(t_i) &= 0, \\ \delta q_-^i(t_f) &= 0; \end{aligned}$$

see Fig. 1 (right).

Integration by parts, using the Keldysh coordinates, and expanding the non-conservative forces in terms of the Keldysh coordinates, Eq. 49 becomes

$$\begin{aligned} 0 = \delta S = \vec{p}_0 \cdot \delta\vec{q}_-(t_i) + \int dt \left\{ \frac{\partial L_1}{\partial q^i} \delta q_1^i + \frac{\partial L_1}{\partial \dot{q}^i} \delta \dot{q}_1^i - \left[\frac{\partial L_2}{\partial q^i} \delta q_2^i + \frac{\partial L_2}{\partial \dot{q}^i} \delta \dot{q}_2^i \right] \right. \\ \left. + \sum_{n=0}^1 \sum_{m=0}^{\infty} \frac{1}{m!} \frac{\partial^m Q_i^{\text{NC}}}{\partial (d^n q_+^j / dt^n)^m} \left(\frac{1}{2} \frac{d^n q_-^j}{dt^n} \right)^m \left[\frac{1 + (-1)^m}{2} \delta q_-^i + \frac{1 - (-1)^m}{2} \delta q_+^i \right] \right\}. \quad (55) \end{aligned}$$

Notice first that the terms in Eq. 49 coming from the conditions Eq. 44b and Eq. 44d have exactly cancelled the boundary terms from the integration by parts. Notice second that the sums in Eq. 55 are both at least linear in $\vec{q}_-$. If we assume for the moment that the equations of motion Eq. 42 subject to the conditions Eq. 44 give the solution $\vec{q}_-(t) = \vec{0}$ for all time, we may drop the terms in Eq. 55 that are linear or higher in $\vec{q}_-(t)$ and its derivatives. We may then straightforwardly see that Eq. 55, with the terms linear or higher in $\vec{q}_-(t)$ and its derivatives dropped, is the stationary point of the action

$$S = \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ L[\vec{q}_1(t), \dot{\vec{q}}_1(t); t] - L[\vec{q}_2(t), \dot{\vec{q}}_2(t); t] + \vec{q}_-(t) \cdot \vec{Q}^{\text{NC}}[\vec{q}_+(t), \dot{\vec{q}}_+(t); t] \right\} \quad (56)$$

for paths $\vec{q}_{1,2}(t)$ allowed to vary but subject to the conditions Eq. 44a and Eq. 44c,

$$\frac{1}{2} [\vec{q}_1(t_i) + \vec{q}_2(t_i)] = \vec{q}_0, \quad (57a)$$

$$\vec{q}_1(t_f) = \vec{q}_2(t_f). \quad (57b)$$

Notice that the variational problem “knows” about the initial conditions from Eq. 25 through the path restriction Eq. 57a and, crucially, through the injection of the initial momentum, effected by the $\vec{p}_0 \cdot \vec{q}_-(t_i)$ term in Eq. 56. While the $\vec{p}_0 \cdot \vec{q}_-(t_i)$ contribution to the action may appear strange at first sight, and also may appear strange in comparison to Hamilton’s original action, this contribution arises naturally from the momentum of the initial quantum wavepacket in the $\hbar \rightarrow 0$ limit of the expectation value of operators such as position or momentum in quantum mechanics [23].

In Eq. 56 and Eq. 57, we have for notational simplicity mixed the $1/2$ and $+/-$ path systems. The more natural way to write Eq. 56 and Eq. 57 is fully in the $+/-$ paths. The conditions Eq. 57 become

$$\vec{q}_+(t_i) = \vec{q}_0, \quad (58a)$$

$$\vec{q}_-(t_f) = \vec{0}, \quad (58b)$$

and the action becomes

$$\begin{aligned} S = \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ L \left[\vec{q}_+(t) + \frac{1}{2} \vec{q}_-(t), \dot{\vec{q}}_+(t) + \frac{1}{2} \dot{\vec{q}}_-(t); t \right] \right. \\ \left. - L \left[\vec{q}_+(t) - \frac{1}{2} \vec{q}_-(t), \dot{\vec{q}}_+(t) - \frac{1}{2} \dot{\vec{q}}_-(t); t \right] + \vec{q}_-(t) \cdot \vec{Q}^{\text{NC}}[\vec{q}_+(t), \dot{\vec{q}}_+(t); t] \right\} \\ = \int_{t_i}^{t_f} dt q_-^i(t) \left[-\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_+^i} + \frac{\partial L}{\partial q_+^i} + Q_i^{\text{NC}}[\vec{q}_+, \dot{\vec{q}}_+; t] + \left(p_{0,i} - \frac{\partial L}{\partial \dot{q}_+^i} \right) \delta(t - t_i) \right]. \quad (59) \end{aligned}$$

To arrive at Eq. 59, we Taylor expanded the first line in powers of $q_-^i(t)$ and $\dot{q}_-^i(t)$ and dropped terms higher than linear in the minus path and its time derivatives, in anticipation that the variation of the action will yield $\vec{q}_-(t) = \vec{0}$ identically. We then integrated by parts; the boundary term evaluated at t_f is zero by the condition $\vec{q}_-(t_f) = \vec{0}$.

III. CONFIRMATION OF THE ACTION

We made a strong assumption in deriving Eq. 56 and Eq. 59: we assumed that, when extremized subject to Eq. 57 and Eq. 58, respectively, the resulting initial conditions and equations of motion would have the unique solution $\vec{q}_-(t) = \vec{0}$ for all time. Let us therefore check that, self-consistently, extremizing Eq. 56 subject to Eq. 57 yields initial conditions and equations of motion whose unique solutions are $\vec{q}_-(t) = \vec{0}$ and $\vec{q}_+(t)$ that solves Eq. 25. (We will check in a moment that Eq. 59 extremized subject to Eq. 58 yields initial conditions and equations of motion whose unique solutions are $\vec{q}_-(t) = \vec{0}$ and $\vec{q}_+(t)$ that solves Eq. 25.)

Replacing $\vec{q}_1(t) = \vec{q}_+(t) + \frac{1}{2}\vec{q}_-(t)$ and $\vec{q}_2(t) = \vec{q}_+(t) - \frac{1}{2}\vec{q}_-(t)$ in Eq. 56 and varying, one finds

$$\begin{aligned}
\delta S &= p_{0,i} \delta q_-^i(t_i) + \int_{t_i}^{t_f} dt \left\{ \left. \frac{\partial L}{\partial q^i} \right|_+ \left(\delta q_+^i + \frac{1}{2} \delta q_-^i \right) + \left. \frac{\partial L}{\partial \dot{q}^i} \right|_+ \frac{d}{dt} \left(\delta q_+^i + \frac{1}{2} \delta q_-^i \right) \right. \\
&\quad \left. - \left[\left. \frac{\partial L}{\partial q^i} \right|_- \left(\delta q_+^i - \frac{1}{2} \delta q_-^i \right) + \left. \frac{\partial L}{\partial \dot{q}^i} \right|_- \frac{d}{dt} \left(\delta q_+^i - \frac{1}{2} \delta q_-^i \right) \right] \right. \\
&\quad \left. + Q_i^{\text{NC}} \delta q_-^i + q_-^j \frac{\partial Q_j^{\text{NC}}}{\partial q^j} \delta q_+^i + q_-^j \frac{\partial Q_j^{\text{NC}}}{\partial \dot{q}^j} \frac{d}{dt} \delta q_+^i \right\} \\
&= \underbrace{\frac{1}{2} \left(\left. \frac{\partial L}{\partial \dot{q}^i} \right|_+ + \left. \frac{\partial L}{\partial \dot{q}^i} \right|_- \right) \delta q_-^i(t_f)}_{=0} \\
&\quad + \left[p_{0,i} - \frac{1}{2} \left(\left. \frac{\partial L}{\partial \dot{q}^i} \right|_+ + \left. \frac{\partial L}{\partial \dot{q}^i} \right|_- \right) \right] \delta q_-^i(t_i) \\
&\quad + \left[\left. \frac{\partial L}{\partial \dot{q}^i} \right|_+ - \left. \frac{\partial L}{\partial \dot{q}^i} \right|_- + \underbrace{q_-^j \frac{\partial Q_j^{\text{NC}}}{\partial \dot{q}^i}}_{=0} \right]_{t_f} \delta q_+^i(t_f) \\
&\quad - \underbrace{\left[\left. \frac{\partial L}{\partial \dot{q}^i} \right|_+ - \left. \frac{\partial L}{\partial \dot{q}^i} \right|_- + q_-^j \frac{\partial Q_j^{\text{NC}}}{\partial \dot{q}^i} \right]_{t_i}}_{=0} \delta q_+^i(t_i) \\
&\quad + \int_{t_i}^{t_f} dt \left\{ \left[\frac{1}{2} \left(\left. \frac{\partial L}{\partial q^i} \right|_+ + \left. \frac{\partial L}{\partial q^i} \right|_- \right) - \frac{1}{2} \frac{d}{dt} \left(\left. \frac{\partial L}{\partial \dot{q}^i} \right|_+ + \left. \frac{\partial L}{\partial \dot{q}^i} \right|_- \right) + Q_i^{\text{NC}} \right] \delta q_-^i \right. \\
&\quad \left. + \left[\left. \frac{\partial L}{\partial q^i} \right|_+ - \left. \frac{\partial L}{\partial q^i} \right|_- - \frac{d}{dt} \left(\left. \frac{\partial L}{\partial \dot{q}^i} \right|_+ - \left. \frac{\partial L}{\partial \dot{q}^i} \right|_- \right) \right. \right. \\
&\quad \left. \left. + q_-^j \frac{\partial Q_j^{\text{NC}}}{\partial q^i} - \frac{d}{dt} \left(q_-^j \frac{\partial Q_j^{\text{NC}}}{\partial \dot{q}^i} \right) \right] \delta q_+^i \right\}, \tag{60}
\end{aligned}$$

where

$$\left. \frac{\partial L}{\partial q^i} \right|_{\pm} \equiv \left. \frac{\partial L}{\partial q^i} \right|_{\substack{q^i(t)=q_+^i(t) \pm \frac{1}{2} q_-^i(t) \\ \dot{q}^i(t)=\dot{q}_+^i(t) \pm \frac{1}{2} \dot{q}_-^i(t)}}, \quad \left. \frac{\partial L}{\partial \dot{q}^i} \right|_{\pm} \equiv \left. \frac{\partial L}{\partial \dot{q}^i} \right|_{\substack{q^i(t)=q_+^i(t) \pm \frac{1}{2} q_-^i(t) \\ \dot{q}^i(t)=\dot{q}_+^i(t) \pm \frac{1}{2} \dot{q}_-^i(t)}}. \tag{61}$$

In Eq. 60, the first term is zero by $\delta q_-^i(t_f) = 0$; similarly, the last term in the square brackets of the third line is zero by $q_-^i(t_f) = 0$. The fourth line is zero by $\delta q_+^i(t_i) = 0$. Since the variations, other than for $\delta q_+^i(t_i) = \delta q_-^i(t_f) = 0$, are independent and arbitrary, by the Fundamental Theorem of Variational Calculus and from $\vec{q}_+(t_i) = \vec{q}_0$ and $\vec{q}_-(t_f) = \vec{0}$,

we must have from Eq. 60 that

$$\vec{q}_-(t_f) = \vec{0}, \quad (62a)$$

$$\left(\frac{\partial L}{\partial \dot{q}^i} \Big|_+ - \frac{\partial L}{\partial \dot{q}^i} \Big|_- \right) \Big|_{t_f} = 0, \quad (62b)$$

$$\frac{\partial L}{\partial q^i} \Big|_+ - \frac{\partial L}{\partial q^i} \Big|_- - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \Big|_+ - \frac{\partial L}{\partial \dot{q}^i} \Big|_- \right) + q_-^j \frac{\partial Q_i^{\text{NC}}}{\partial q^j} - \frac{d}{dt} \left(q_-^j \frac{\partial Q_i^{\text{NC}}}{\partial \dot{q}^j} \right) = 0, \quad (62c)$$

$$\vec{q}_+(t_i) = \vec{q}_0, \quad (62d)$$

$$\frac{1}{2} \left(\frac{\partial L}{\partial \dot{q}^i} \Big|_+ + \frac{\partial L}{\partial \dot{q}^i} \Big|_- \right) \Big|_{t_i} = p_{0,i}, \quad (62e)$$

$$\frac{1}{2} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}^i} \Big|_+ + \frac{\partial L}{\partial \dot{q}^i} \Big|_- \right) - \frac{1}{2} \left(\frac{\partial L}{\partial q^i} \Big|_+ + \frac{\partial L}{\partial q^i} \Big|_- \right) = Q_i^{\text{NC}}. \quad (62f)$$

One sees that the two conditions on the paths (and their variations) have, by extremizing the action Eq. 56, yielded four initial conditions, Eq. 62a, Eq. 62b, Eq. 62d, and Eq. 62e, and two equations of motion, Eq. 62c and Eq. 62f. For the minus path, the initial conditions are set at the final time, and the equations of motion propagate those conditions backwards in time; see Fig. 1 (right). Considering the Taylor expansion of Eq. 62b in minus paths, one can see that Eq. 62b is, given the form of L given in Eq. 36, equivalent to $\vec{q}_-(t_f) = \vec{0}$. Notice that $\vec{q}_-(t) = \vec{0}$ is a solution to Eq. 62a–62c; by the Picard–Lindelöf theorem, $\vec{q}_-(t) = \vec{0}$ is the unique solution to Eq. 62a–62c.

With $\vec{q}_-(t) = \vec{0}$, Eq. 62d–62f become identical to Eq. 38, which itself is identical to Eq. 25. We have thus shown that extremization of the action Eq. 56 subject to paths satisfying Eq. 57 is exactly equivalent to Newton’s Second Law (for any forces that depend, at most, up to the first time derivative of the path).

Let us now check that the extremization of the action Eq. 59 subject to paths that satisfy Eq. 58 is equivalent to Eq. 25. Extremizing Eq. 59 subject to paths that satisfy Eq. 58 yields, after careful consideration of the variations and integrations by parts,

$$\begin{aligned} \delta S = & \left(p_{0,i} - \frac{\partial L}{\partial \dot{q}_+^i} \right) \delta q_-^i(t_i) + \dot{q}_-^j \frac{\partial^2 L}{\partial \dot{q}_+^i \partial \dot{q}_+^j} \delta q_+^i(t_f) + \int_{t_i}^{t_f} dt \left\{ \delta q_-^i(t) \left(-\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_+^i} + \frac{\partial L}{\partial q_+^i} + Q_i^{\text{NC}}(\vec{q}_+, \dot{\vec{q}}_+; t) \right) \right. \\ & \left. + \delta q_+^i(t) \left(\dot{q}_-^j \frac{\partial^2 L}{\partial \dot{q}_+^i \partial \dot{q}_+^j} + q_-^j \left(\frac{\partial^2 L}{\partial q_+^i \partial q_+^j} + \frac{\partial Q_j^{\text{NC}}}{\partial q_+^i} \right) - \frac{d}{dt} \left(\dot{q}_-^j \frac{\partial^2 L}{\partial \dot{q}_+^i \partial \dot{q}_+^j} + q_-^j \frac{\partial Q_j^{\text{NC}}}{\partial \dot{q}_+^i} \right) \right) \right\}. \end{aligned} \quad (63)$$

Since we restrict our analysis to generalized coordinates such that $\partial^2 L / \partial \dot{q}^i \partial \dot{q}^j$ is non-degenerate, one sees from Eq. 63 that setting $\delta S = 0$ implies

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_+^i} - \frac{\partial L}{\partial q_+^i} &= Q_i^{\text{NC}}(\vec{q}_+, \dot{\vec{q}}_+; t), \\ \vec{q}_+(t_i) &= \vec{q}_0, \\ \frac{\partial L}{\partial \dot{q}_+^i} \Big|_{t_i} &= p_{0,i}, \end{aligned} \quad (64)$$

and

$$\begin{aligned} \dot{q}_-^j \left(\frac{\partial^2 L}{\partial \dot{q}_+^i \partial \dot{q}_+^j} + q_-^j \frac{\partial Q_j^{\text{NC}}}{\partial q_+^i} \right) &= \dot{q}_-^j \frac{\partial^2 L}{\partial \dot{q}_+^i \partial \dot{q}_+^j} + q_-^j \left(\frac{\partial^2 L}{\partial q_+^i \partial q_+^j} + \frac{\partial Q_j^{\text{NC}}}{\partial q_+^i} \right), \\ \vec{q}_-(t_i) &= \vec{0}, \\ \dot{\vec{q}}_-(t_i) &= \vec{0}. \end{aligned} \quad (65)$$

Clearly $\vec{q}_-(t) = \vec{0}$ is a solution to Eq. 65; by the Picard–Lindelöf theorem, $\vec{q}_-(t) = \vec{0}$ is the unique solution. Our action Eq. 59 subject to paths Eq. 58 thus self-consistently yields $\vec{q}_-(t) = \vec{0}$ for all time as required by the derivation of Eq. 59. Further, it is clear that Eq. 64 is identical to Eq. 38, which is equivalent to Newton’s Second Law Eq. 25.

IV. PHASE SPACE ACTION AND HAMILTON'S EQUATIONS

The Hamilton's Equations associated with Eq. 25, equivalently Eq. 38, are

$$\begin{aligned}\frac{dp_i}{dt} &= -\frac{\partial H}{\partial q^i} + Q_i^{\text{NC}}, \\ \frac{dq^i}{dt} &= \frac{\partial H}{\partial p_i}, \\ \vec{q}(t_i) &= \vec{q}_0, \\ \vec{p}(t_i) &= \vec{p}_0.\end{aligned}\tag{66}$$

If one defines the Hamiltonian as the usual Legendre transform of the Lagrangian,

$$H[\vec{p}_a, \vec{q}_a; t] \equiv p_{a,i} \dot{q}_a^i - L[\vec{q}_a, \dot{\vec{q}}_a; t], \quad a = 1, 2,\tag{67}$$

where

$$p_{a,i} \equiv \frac{\partial L}{\partial \dot{q}_a^i},\tag{68}$$

then the phase space action equivalent to the configuration space action Eq. 56 is

$$S = \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt' \left\{ \vec{p}_1 \cdot \dot{\vec{q}}_1 - H[\vec{p}_1, \vec{q}_1; t'] - (\vec{p}_2 \cdot \dot{\vec{q}}_2 - H[\vec{p}_2, \vec{q}_2; t']) + \vec{q}_- \cdot \vec{Q}^{\text{NC}}[\vec{q}_+, \vec{p}_+; t'] \right\}.\tag{69}$$

Then, extremizing Eq. 69 with respect to $\vec{q}_\pm(t)$ and $\vec{p}_\pm(t)$ subject to

$$\begin{aligned}\frac{1}{2}[\vec{q}_1(t_i) + \vec{q}_2(t_i)] &= \vec{q}_0, \\ \vec{q}_1(t_f) &= \vec{q}_2(t_f),\end{aligned}\tag{70}$$

yields the initial conditions

$$\begin{aligned}\vec{q}_+(t_i) &= \vec{q}_0, \\ \vec{p}_+(t_i) &= \vec{p}_0, \\ \vec{q}_-(t_f) &= \vec{0}, \\ \vec{p}_-(t_f) &= \vec{0},\end{aligned}\tag{71}$$

and equations of motion for the minus path and momentum whose unique solution is

$$\begin{aligned}\vec{q}_-(t) &= \vec{0}, \\ \vec{p}_-(t) &= \vec{0},\end{aligned}\tag{72}$$

and the equations of motion for the plus path and momentum of

$$\begin{aligned}\frac{dp_{+,i}}{dt} &= -\frac{\partial H}{\partial q_+^i} + Q_i^{\text{NC}}[\vec{q}_+, \vec{p}_+; t], \\ \frac{dq_+^i}{dt} &= \frac{\partial H}{\partial p_{+,i}},\end{aligned}\tag{73}$$

which are the classical Hamilton's Equations Eq. 66.

In the same way as for the configuration space action, we may write our phase space action for Eq. 69 fully in the Keldysh coordinates, Taylor expand in the minus coordinates, and truncate under the assumption that $\vec{q}_-(t) = \vec{p}_-(t) = \vec{0}$. One then finds the action

$$S = [\vec{p}_0 - \vec{p}_+(t_i)] \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ p_{-,i} \left[\dot{q}_+^i - \frac{\partial H}{\partial p_{+,i}} \right] - q_-^i \left[\dot{p}_{+,i} + \frac{\partial H}{\partial q_+^i} - Q_i^{\text{NC}}[\vec{q}_+, \vec{p}_+; t] \right] \right\}.\tag{74}$$

Extremizing this action Eq. 74 with respect to $\vec{q}_\pm(t)$ and $\vec{p}_\pm(t)$ subject to

$$\begin{aligned}\vec{q}_+(t_i) &= \vec{q}_0, \\ \vec{q}_-(t_f) &= \vec{0}\end{aligned}\tag{75}$$

yields again Eq. 71 and 73.

We have thus shown how to properly take the Legendre transform of our Lagrangian to yield an action whose extremization yields the classical path that solves the initial value problem of Newton's Second Law for any force (that depends up to the first time derivative of the path). Note that the definition of momentum used here, Eq. 68, differs from Galley [9, 27], who includes the non-conservative force in the definition of momentum.

V. GENERALIZATION TO ANY FORCES

Let us now generalize to the situation in which the force depends on any number of time derivatives of the path. Such a force is not fundamental. Nevertheless, there are approximations that lead to forces that depend on higher derivatives of the path, such as the Abraham–Lorentz force [28], which depends on the third time derivative of the path. Alternatively, one might consider, e.g., an external feedback force tuned to depend on an arbitrary function of the various time derivatives of the path, for example a force feedback dependent on the acceleration of the particle [29]. Newton's Second Law applies to all such physical setups. Let us now show how to incorporate any such setup in a variational principle.

The key difficulty to constructing an action similar to Eq. 56 or Eq. 59 whose variation yields Newton's Second Law is the setting of the initial conditions for the higher time derivatives of the path. Unfortunately, we do not currently have some $\hbar \rightarrow 0$ calculation of some quantum system to provide guidance. Naively including the higher time derivatives of the path in $\vec{Q}^{\text{NC}}$ in the action does not lead to the correct initial conditions as the integration by parts does not inject the initial conditions for these higher time derivatives of the path. It also does not seem like one can construct some term similar to the $\vec{p}_0 \cdot \vec{q}_-(t_i)$ that cancels the boundary terms from the integration by parts that injects the initial conditions for these higher time derivative terms. For example,

$$q_-^i(t_i) \frac{\partial Q_{0,i}^{\text{NC}}}{\partial (d^I q_+^j / dt^I)} \left(\frac{d^{I-1} q_+^j}{dt^{I-1}} \Big|_{t_i} - \frac{d^{I-1} q_{0,+}^j}{dt^{I-1}} \right)\tag{76}$$

will cancel a boundary term upon variation but cannot provide the initial value for $d^N q^j / dt^N$, the highest time derivative needed. Eq. 76 is also problematic as it will interfere with the momentum injection from the $\vec{p}_0 \cdot \vec{q}_-(t_i)$ term. One may alternatively try to utilize Ostrogradsky's construction [30]. However, Newton's Second Law treats only the path and its first time derivative as dynamical; the Lagrangian only depends on $\vec{q}$ and $\dot{\vec{q}}$. As a result, the higher momenta from Ostrogradsky will come solely from the dependence of $\vec{Q}^{\text{NC}}$ on higher time derivatives of the path, generating unwanted contributions to the equations of motion.

The route we will take for constructing an action whose variation yields Newton's Second Law for forces that depend on any orders of the time derivative of the path is similar to that done in [23]: we will introduce new coordinates $\vec{q}^{(I)}(t)$, connected to the I^{th} time derivative of the path through Lagrange multipliers. We may then consider variations of the $\vec{q}^{(I)}$ degrees of freedom as independent of the variations of the path. In particular, we may consider $\delta \vec{q}^{(I)}(t)$ and not need to integrate by parts. Further, we are now free to set $\vec{q}^{(I)}(t_i) = d^I \vec{q} / dt^I|_{t_i}$; as a result, $\delta \vec{q}^{(I)}(t)$ is independent and arbitrary, except that $\delta \vec{q}^{(I)}(t_i) = \vec{0}$.

With the above as context, we proceed as follows. We start with

$$\vec{F}_1 = m \ddot{\vec{x}}_1 \quad \text{and} \quad \vec{F}_2 = m \ddot{\vec{x}}_2,\tag{77}$$

where $\vec{F}_i = \vec{F}(\{d^I \vec{x}_i / dt^I\}_{I=0}^N; t)$, subject to the initial conditions

$$\begin{aligned}\frac{1}{2} \left(\frac{d^I \vec{x}_1}{dt^I} + \frac{d^I \vec{x}_2}{dt^I} \right) \Big|_{t_i} &= \frac{d^I \vec{x}}{dt^I} \Big|_{t_i}, \quad I = 0, \dots, N, \\ \vec{x}_1(t_f) - \vec{x}_2(t_f) &= \vec{0}, \\ \dot{\vec{x}}_1(t_f) - \dot{\vec{x}}_2(t_f) &= \vec{0}.\end{aligned}\tag{78}$$

Note that, while we need $N + 1$ conditions at t_i , we continue to only need two conditions at t_f . We transition to generalized Keldysh coordinates; we now have

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_a^i} - \frac{\partial L}{\partial q_a^i} = Q_{a,i}^{\text{NC}} \left[\{\vec{q}_a^{(I)}(t)\}_{I=0}^N; t \right], \quad a = 1, 2, \quad (79a)$$

$$\vec{q}_+^{(I)}(t) = \frac{d^I \vec{q}_+}{dt^I}, \quad I = 2, \dots, N, \quad (79b)$$

$$\frac{1}{2} [\vec{q}_1(t_i) + \vec{q}_2(t_i)] = \vec{q}_0, \quad (79c)$$

$$\frac{1}{2} \left[\frac{\partial L}{\partial \dot{q}_1^i} \Big|_{t_i} + \frac{\partial L}{\partial \dot{q}_2^i} \Big|_{t_i} \right] = p_{0,i}, \quad (79d)$$

$$\vec{q}_+^{(I)}(t_i) = \frac{d^I \vec{q}_+}{dt^I} \Big|_{t_i}, \quad I = 2, \dots, N, \quad (79e)$$

$$\vec{q}_1(t_f) - \vec{q}_2(t_f) = \vec{0}, \quad (79f)$$

$$\frac{\partial L}{\partial \dot{q}_1^i} \Big|_{t_f} - \frac{\partial L}{\partial \dot{q}_2^i} \Big|_{t_f} = 0. \quad (79g)$$

We may proceed exactly as was done in Eq. 49, but now we also add the zeros

$$\int_{t_i}^{t_f} dt \delta \vec{\lambda}_-^{(I)}(t) \cdot \left(\frac{d^I \vec{q}_+}{dt^I} - \vec{q}_+^{(I)}(t) \right),$$

where $\delta \vec{\lambda}_-^{(I)}(t)$ are independent and arbitrary; there are no restrictions on the $\delta \vec{\lambda}_-^{(I)}(t)$.

Under the assumption that, after variation, $\vec{\lambda}_-^{(I)}(t)$ will be set to $\vec{0}$, we arrive at the following two actions

$$S = \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ L[\vec{q}_1, \dot{\vec{q}}_1; t] - L[\vec{q}_2, \dot{\vec{q}}_2; t] + \sum_{I=2}^N \vec{\lambda}_-^{(I)} \cdot \left(\frac{d^I \vec{q}_+}{dt^I} - \vec{q}_+^{(I)} \right) + \vec{q}_- \cdot \vec{Q}^{\text{NC}}(\vec{q}_+, \dot{\vec{q}}_+, \{\vec{q}_+^{(I)}\}_{I=2}^N; t) \right\} \quad (80)$$

and

$$S = \left(p_{0,i} - \frac{\partial L}{\partial \dot{q}_+^i} \right) q_-^i(t_i) + \int_{t_i}^{t_f} dt \left\{ q_-^i(t) \left(-\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_+^i} + \frac{\partial L}{\partial q_+^i} + Q_i^{\text{NC}}(\vec{q}_+, \dot{\vec{q}}_+, \{\vec{q}_+^{(I)}\}_{I=2}^N; t) \right) + \sum_{I=2}^N \vec{\lambda}_-^{(I)} \cdot \left(\frac{d^I \vec{q}_+}{dt^I} - \vec{q}_+^{(I)} \right) \right\}. \quad (81)$$

One can see relatively straightforwardly that Eq. 80 or Eq. 81, under variations subject to Eq. 79c–79g, yield initial conditions and equations of motion whose solution are

$$\vec{q}_-(t) = \vec{0}, \quad (82a)$$

$$\vec{\lambda}_-^{(I)}(t) = \vec{0}, \quad I = 2, \dots, N; \quad (82b)$$

i.e., we see that, as for our previous actions, the extremization of Eq. 80 or Eq. 81 subject to Eq. 79c–79g self-consistently yields all minus coordinates identically 0 for all time. Inserting Eq. 82a and 82b in the equations of motion for the plus coordinates yields

$$\begin{aligned} \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} &= Q_i^{\text{NC}} \left(\left\{ \frac{d^I q^i}{dt^I} \right\}_{I=0}^N \right), \\ \frac{d^I \vec{q}}{dt^I} \Big|_{t_i} &= \frac{d^I \vec{q}_0}{dt^I}, \quad I = 0, \dots, N. \end{aligned} \quad (83)$$

Since Eq. 83 is equivalent to Newton's Second Law, Eq. 25, we have thus succeeded in formulating an action principle in configuration space for any force. When moving to phase space, the time derivatives of the path of second order and higher are to be considered non-dynamical; therefore one keeps the definition of 1/2 momenta, Eq. 68, and one writes the non-conservative force in terms of $\vec{q}_+(t)$, $\vec{p}_+(t)$, and the new coordinates $\vec{q}^{(I)}(t)$,

$$\vec{Q}^{\text{NC}}[\vec{p}_+, \vec{q}_+, \{\vec{q}^{(I)}\}_{I=2}^N; t]. \quad (84)$$

VI. FORCES OF CONSTRAINT

From the work of [4], it is clear that the equations of motion for a particle subject to $k = 1, \dots, m$ constraints

$$h_k \left(\left\{ \frac{d^I q^i}{dt^I} \right\}_{I=0}^M; t \right) = 0,$$

as a function of generalised coordinates $q^i(t)$ and their time derivatives, that do no virtual work,

$$\frac{\partial h_k}{\partial \frac{d^M q^i}{dt^M}} \delta q^i = 0,$$

satisfy the equations of motion

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} - \frac{\partial L}{\partial q^i} = Q_i^{\text{NC}} + \lambda^k \frac{\partial h_k}{\partial \frac{d^M q^i}{dt^M}}, \quad (85)$$

where the λ_i^k are Lagrange multipliers.

If we, as is now usual, consider a doubling of the degrees of freedom, we then have that

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_a^i} - \frac{\partial L}{\partial q_a^i} = Q_{a,i}^{\text{NC}} [\{\vec{q}_a^{(I)}(t)\}_{I=0}^N] + \lambda_a^k(t) \frac{\partial h_k}{\partial \frac{d^M q^i}{dt^M}}, \quad a = 1, 2, \quad (86a)$$

$$\vec{q}_+^{(I)}(t) = \frac{d^I \vec{q}_+}{dt^I}, \quad I = 2, \dots, \max(M, N), \quad (86b)$$

$$h_k \left(\left\{ \frac{d^I q_a^i}{dt^I} \right\}_{I=0}^M; t \right) = 0, \quad a = 1, 2, \quad (86c)$$

$$\frac{1}{2} [\vec{q}_1(t_i) + \vec{q}_2(t_i)] = \vec{q}_0, \quad (86d)$$

$$\frac{1}{2} \left\{ \frac{\partial L}{\partial \dot{q}_1^i} \Big|_{t_i} + \frac{\partial L}{\partial \dot{q}_2^i} \Big|_{t_i} \right\} = p_{0,i}, \quad (86e)$$

$$\vec{q}_+^{(I)}(t_i) = \frac{d^I \vec{q}_+}{dt^I} \Big|_{t_i}, \quad I = 2, \dots, \max(M, N), \quad (86f)$$

$$\vec{q}_1(t_f) - \vec{q}_2(t_f) = \vec{0}, \quad (86g)$$

$$\frac{\partial L}{\partial \dot{q}_1^i} \Big|_{t_f} - \frac{\partial L}{\partial \dot{q}_2^i} \Big|_{t_f} = 0. \quad (86h)$$

Following exactly the same logic as in the previous sections, we may construct a $\delta S = 0$ by adding in the usual way the zeroes from the previous subequations and see that, under the assumption that all minus coordinates are zero for all time, δS is the variation of

$$\begin{aligned} S = \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ L[\vec{q}_1, \dot{\vec{q}}_1; t] - L[\vec{q}_2, \dot{\vec{q}}_2; t] + q_-^i(t) \left(Q_i^{\text{NC}} \left[\left\{ \frac{d^I \vec{q}_+}{dt^I} \right\}_{I=0}^N; t \right] + \lambda_+^k(t) \frac{\partial h_k}{\partial \frac{d^M q^i}{dt^M}} \right) \right. \\ \left. + \lambda_-^k(t) h_k \left[\left\{ \frac{d^I \vec{q}_+}{dt^I} \right\}_{I=0}^M; t \right] + \sum_{I=2}^{\max(M, N)} \vec{\lambda}_-^{(I)}(t) \cdot \left(\frac{d^I \vec{q}_+}{dt^I} - \vec{q}_+^{(I)}(t) \right) \right\}, \end{aligned} \quad (87)$$

or, equivalently,

$$\begin{aligned} S = \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ q^i(t) \cdot \left(-\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_+^i} + \frac{\partial L}{\partial q_+^i} + Q_i^{\text{NC}} \left[\left\{ \frac{d^I \vec{q}_+}{dt^I} \right\}_{I=0}^N; t \right] + \lambda_+^k(t) \frac{\partial h_k}{\partial \frac{d^M q^i}{dt^M}} \right) \right. \\ \left. + \lambda_-^k(t) h_k \left[\left\{ \frac{d^I \vec{q}_+}{dt^I} \right\}_{I=0}^M; t \right] + \sum_{I=2}^{\max(M, N)} \vec{\lambda}_-^{(I)} \cdot \left(\frac{d^I \vec{q}_+}{dt^I} - \vec{q}_+^{(I)} \right) \right\}. \end{aligned} \quad (88)$$

In phase space, the correct action to vary is

$$\begin{aligned}
S = \vec{p}_0 \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ (\vec{p}_1 \cdot \dot{\vec{q}}_1 - H[\vec{p}_1, \vec{q}_1; t]) - (\vec{p}_2 \cdot \dot{\vec{q}}_2 - H[\vec{p}_2, \vec{q}_2; t]) \right. \\
\left. + q_-^i(t) \cdot \left(Q_i^{\text{NC}} \left[\vec{q}_+, \vec{p}_+, \left\{ \vec{q}_+^{(I)} \right\}_{I=2}^N; t \right] + \lambda_+^k(t) \frac{\partial h_k}{\partial \frac{d^M q^i}{dt^M}} \right) \right. \\
\left. + \lambda_-^k(t) h_k \left[\vec{p}_+, \vec{q}_+, \left\{ \frac{d^I \vec{q}_+}{dt^I} \right\}_{I=2}^N; t \right] + \sum_{I=2}^{\max(M,N)} \vec{\lambda}_-^{(I)} \cdot \left(\frac{d^I \vec{q}_+}{dt^I} - \vec{q}_+^{(I)} \right) \right\}, \quad (89)
\end{aligned}$$

or, equivalently,

$$\begin{aligned}
S = [\vec{p}_0 - \vec{p}_+(t_i)] \cdot \vec{q}_-(t_i) + \int_{t_i}^{t_f} dt \left\{ p_{-,i} \cdot \left[\dot{q}_+^i - \frac{\partial H}{\partial p_{+,i}} \right] \right. \\
\left. - q_-^i \cdot \left[\dot{p}_{+,i} + \frac{\partial H}{\partial q_+^i} - Q_i^{\text{NC}} \left[\vec{q}_+, \vec{p}_+, \left\{ \vec{q}_+^{(I)} \right\}_{I=2}^N; t \right] + \lambda_+^k(t) \frac{\partial h_k}{\partial \frac{d^M q^i}{dt^M}} \right] \right. \\
\left. + \lambda_-^k(t) h_k \left[\vec{p}_+, \vec{q}_+, \left\{ \frac{d^I \vec{q}_+}{dt^I} \right\}_{I=2}^N; t \right] + \sum_{I=2}^{\max(M,N)} \vec{\lambda}_-^{(I)} \cdot \left(\frac{d^I \vec{q}_+}{dt^I} - \vec{q}_+^{(I)} \right) \right\}. \quad (90)
\end{aligned}$$

VII. CONCLUSIONS

We derived from first principles in classical mechanics (i.e. directly from Newton's Second Law as an initial value problem) configuration space and related Legendre-transformed phase space actions whose extremization is equivalent to Newton's Second Law as an initial value problem for any set of forces: conservative forces, forces of constraint that do no virtual work, and general non-conservative forces. The extremization procedure requires only knowledge of the initial conditions of the problem, unlike the usual Hamilton's Principle that requires the acausal knowledge of the final position of the particle during the extremization procedure. We provided the exact construction of the configuration and phase space actions for any set of conservative, constraint (that do no virtual work), and general non-conservative forces.

Similar to the Schwinger–Keldysh path integral formulation of the expectation value of operators in quantum mechanics, our action doubles the degrees of freedom to two paths in generalized coordinates. The extremization of any of the actions returns two sets of two initial conditions and two equations of motion. For the configuration space actions, extremization yields 1) an auxiliary path, $\vec{q}_-(t)$, with initial conditions at the final time of $\vec{q}_-(t_f) = \vec{0}$ and $\dot{\vec{q}}_-(t_f) = \vec{0}$ and equations of motion that propagate those initial conditions backwards in time with the unique solution $\vec{q}_-(t) = \vec{0}$ for all time and 2) the physical path, $\vec{q}_+(t)$, whose initial conditions are the initial conditions of the physical setup, $\{d^I \vec{q}_+/dt^I|_{t_i}\}_{I=0}^{\max(1,M,N)}$, where M is the largest time derivative of the path that any constraint depends on and N is the largest time derivative of the path that any applied force depends on, and which satisfies the Euler–Lagrange equations of motion for any general forces; i.e. the physical path solves Newton's Second Law as an initial value problem for any general forces. For the phase space actions, extremization yields 1) initial conditions for the auxiliary path $\vec{q}_-(t)$ and its associated auxiliary momentum $\vec{p}_-(t)$, $\vec{q}_-(t_f) = \vec{p}_-(t_f) = \vec{0}$, and equations of motion that propagate those values backwards in time with the unique solution $\vec{q}_-(t) = \vec{p}_-(t) = \vec{0}$ for all time and 2) initial conditions of the physical path $\vec{q}_+(t)$, its associated physical momentum $\vec{p}_+(t)$, and any higher derivatives necessary for the constraints or applied forces to their physical values and yields Hamilton's Equations for any force as the equations of motion; thus the physical path and momentum solve Newton's Second Law as an initial value problem for any general forces. Interestingly, for forces and constraints that depend at most on the first time derivative of the path, the extremization procedure requires only setting $\vec{q}_+(t_i)$ to the initial position $\vec{q}_0$ and $\vec{q}_-(t_f) = \vec{0}$ —i.e. only setting the two variations $\delta \vec{q}_+(t_i)$ and $\delta \vec{q}_-(t_f)$ to zero—to generate the four initial conditions and two equations of motion.

The actions derived here are similar to that postulated by Galley [9] and Bateman [10] but differ in important ways. On a practical level, first, the actions derived here include an additional term $\vec{p}_0 \cdot \vec{q}_-(t_i)$ —necessary for setting the velocity (in configuration space) or momentum (in phase space) initial condition for the physical path—which is absent in Galley's and Bateman's postulated actions. This additional term is consistent with the $\hbar \rightarrow 0$ limit of the quantum mechanics of unconstrained motion subject to conservative forces [23]. Second, in contrast

to [9, 10], we provide an explicit construction of our actions for any given forces. Third, for applied forces that depend on up to at most the first time derivative of the path, the extremization procedure described here requires only $\delta\vec{q}_+(t_i) = \delta\vec{q}_-(t_f) = \vec{0}$, consistent with quantum mechanics [23], but which differs from the requirement of $\delta\vec{q}_+(t_i) = \delta\vec{q}_-(t_i) = \delta\vec{q}_-(t_f) = \vec{0}$ and $\dot{\vec{q}}_-(t_f) = \vec{0}$ in [9]. Numerical extremization of the action Eq. 56 derived here subject only to the two conditions $\delta\vec{q}_+(t_i) = \delta\vec{q}_-(t_f) = \vec{0}$ yielded physical paths consistent with analytic solutions to Newton’s Second Law as an initial value problem [31] whereas numerical extremization of the action of [9] subject to the four conditions $\delta\vec{q}_+(t_i) = \delta\vec{q}_-(t_i) = \delta\vec{q}_-(t_f) = \vec{0}$ and $\dot{\vec{q}}_-(t_f) = \vec{0}$ yielded paths with unphysical jumps at the endpoints, a classic sign of overspecification in numerical analysis. The actions derived here also go beyond [9, 10] by including forces of constraint.

Similarly, our work here modifies the action we postulated in [24] for velocity-dependent non-holonomic constraints: we correct our action with the $\vec{p}_0 \cdot \vec{q}_-(t_i)$ term and generalize to arbitrary forces of constraint that perform no virtual work.

There are many potential directions of future work. While almost all point particle mechanics problems may be formulated as a Newton II-type equation of motion involving forces, many problems are much more naturally treated within an action formulation. The actions here are thus well-suited for, e.g., dissipative systems with non-holonomic constraints. Systems with topological additions to the action, e.g. systems with magnetic monopoles, utilizing Wess-Zumino type terms in the action, etc., and point particles with intrinsic gauge degrees of freedom are much more naturally formulated as action principles. Thus the actions found here are natural for treating dissipative and/or constrained charged point particle processes. While the evolution of charge within group space can in principle be treated by Newton’s Second Law, a point particle with intrinsic spin can *only* be treated approximately by Newton’s Second Law as the limit of an extended rigid body. Approximating a point particle as an extended object will break down when curvature is rapidly changing, and thus the actions here are uniquely suited for treating the dissipative and/or constrained motion of classical point particles with spin in geometries with rapidly changing curvature.

The real power of action principles lies in their treatment of symmetries and in field theories. One would like to understand Noether’s Theorem and the consequences of symmetries in an initial value variational formulation of classical mechanics. Galley and co-authors [27] considered the consequences of Noether using Galley’s original postulated action [9]; one should determine if the injection of the initial momentum and the different variations of the path in this work lead to different results from [27]. The work here also allows for the first time the investigation of the consequences of Noether’s theorem for systems subject to non-holonomic constraints [4, 32–34]. One would further like to extend the formalism developed in [23] and here for classical field theories. There is no obvious obstruction to doing so. Applications in dissipative effective field theories have already been considered using the original Galley postulated action [19–21]; one would like to understand how the modifications to the action and extremization procedure here modify those results. Of extreme interest is the generalization of the actions here to fields with general non-holonomic constraints [35]. It does not seem as though there are physical consequences to using a vakonomic-type approach and simply adjoining a non-holonomic gauge fixing term to the electromagnetic field Lagrangian with a Lagrange multiplier [36]. However, there do appear to be physical consequences for using a vakonomic-type approach to enforce the non-holonomic constraint of the conformal gauge on a relativistic string [31]; a first-principles derivation of the motion of a relativistic string subject to the conformal gauge using a generalization of the actions here would be of great interest.

ACKNOWLEDGEMENTS

WAH gratefully acknowledges invaluable discussions with Klaas Landsman, Alexander Rothkopf, and Ben Bert.

This research was conducted in part by WAH while visiting the Okinawa Institute of Science and Technology (OIST) through the Theoretical Sciences Visiting Program (TSVP). WAH thanks the National Research Foundation and the SA-CERN collaboration for their generous financial support during the course of this work.

-
- [1] I. Newton, *Philosophiæ Naturalis Principia Mathematica* (Royal Society, London, 1687).
 - [2] L. Euler, *Methodus Inveniendi Lineas Curvas Maximi Minimive Proprietate Gaudentes* (Bousquet, Lausanne & Geneva, 1744).
 - [3] J.-L. Lagrange, *Mécanique Analytique* (La Veuve Desaint, Paris, 1788).
 - [4] M. R. Flannery, *Journal of Mathematical Physics* **52**, 032705 (2011).
 - [5] W. R. Hamilton, *Philosophical Transactions of the Royal Society of London* **124**, 247 (1834).
 - [6] W. R. Hamilton, *Philosophical Transactions of the Royal Society of London* **125**, 95 (1835).

- [7] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations* (McGraw-Hill, 1955) 9th reprint (1987).
- [8] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics*, 3rd ed. (Addison-Wesley, San Francisco, 2002).
- [9] C. R. Galley, *Phys. Rev. Lett.* **110**, 174301 (2013), [arXiv:1210.2745 \[gr-qc\]](#).
- [10] H. Bateman, *Physical Review* **38**, 815 (1931).
- [11] G. Mogull, J. Plefka, and J. Steinhoff, *JHEP* **02**, 048 (2021), [arXiv:2010.02865 \[hep-th\]](#).
- [12] O. Birnholtz, S. Hadar, and B. Kol, *Phys. Rev. D* **88**, 104037 (2013), [arXiv:1305.6930 \[hep-th\]](#).
- [13] R. A. Porto, *Phys. Rept.* **633**, 1 (2016), [arXiv:1601.04914 \[hep-th\]](#).
- [14] M. Levi, *Rept. Prog. Phys.* **83**, 075901 (2020), [arXiv:1807.01699 \[hep-th\]](#).
- [15] A. Edison, *JHEP* **02**, 016 (2025), [arXiv:2409.17222 \[hep-th\]](#).
- [16] J. Polonyi, *Annals Phys.* **467**, 169694 (2024), [arXiv:2307.12320 \[quant-ph\]](#).
- [17] F. E. Riewe, *Phys. Rev. E* **55**, 3581 (1997).
- [18] K. H. Belgacem, F. Jiménez, and S. Ober-Blöbaum, *Journal of Nonlinear Science* **35**, 38 (2025).
- [19] S. Endlich, A. Nicolis, R. A. Porto, and J. Wang, *Phys. Rev. D* **88**, 105001 (2013), [arXiv:1211.6461 \[hep-th\]](#).
- [20] M. Crossley, P. Glorioso, and H. Liu, *JHEP* **09**, 095 (2017), [arXiv:1511.03646 \[hep-th\]](#).
- [21] S. Grozdanov and J. Polonyi, *Phys. Rev. D* **91**, 105031 (2015), [arXiv:1305.3670 \[hep-th\]](#).
- [22] A. Rothkopf and J. Nordström, *J. Comput. Phys.* **498**, 112652 (2024), [arXiv:2307.04490 \[math.NA\]](#).
- [23] W. A. Horowitz and A. Rothkopf, (2026), [arXiv:2603.02350 \[physics.class-ph\]](#).
- [24] A. Rothkopf and W. A. Horowitz, *Physical Review E* (2024), 10.1103/rch9-wsbw, [arXiv:2409.11063 \[physics.class-ph\]](#).
- [25] P. A. M. Dirac, *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, Yeshiva University, New York, 1964).
- [26] K. Sundermeyer, *CONSTRAINED DYNAMICS WITH APPLICATIONS TO YANG-MILLS THEORY, GENERAL RELATIVITY, CLASSICAL SPIN, DUAL STRING MODEL*, Vol. 169 (1982).
- [27] C. R. Galley, D. Tsang, and L. C. Stein, (2014), [arXiv:1412.3082 \[math-ph\]](#).
- [28] P. A. M. Dirac, *Proc. Roy. Soc. A* **167**, 148 (1938).
- [29] C. Ott, *Cartesian Impedance Control of Redundant and Flexible-Joint Robots*, Springer Tracts in Advanced Robotics, Vol. 49 (Springer Berlin, Heidelberg, 2008).
- [30] M. Ostrogradsky, *Mémoires de l'Académie Impériale des Sciences de Saint-Pétersbourg Sixième série*, **6**, 385 (1850).
- [31] A. Rothkopf, W. A. Horowitz, and J. Nordström, in preparation (2026).
- [32] M. R. Flannery, *American Journal of Physics* **73**, 265 (2005).
- [33] A. M. Bloch, *Nonholonomic Mechanics and Control*, 2nd ed., Interdisciplinary Applied Mathematics, Vol. 24 (Springer, New York, NY, 2015).
- [34] W. A. Horowitz and A. Rothkopf, (2024), [arXiv:2408.14420 \[math-ph\]](#).
- [35] B. Bert and W. A. Horowitz, (2025), [arXiv:2505.20684 \[hep-th\]](#).
- [36] B. Bert, A. Rothkopf, and W. A. Horowitz, in preparation (2026).