Variational principles for nonlinear PDE systems via duality

Amit Acharya*

Abstract

A formal methodology for developing variational principles corresponding to a given nonlinear PDE system is discussed. The scheme is demonstrated in the context of the incompressible Navier-Stokes equations, systems of first-order conservation laws, and systems of Hamilton-Jacobi equations.

1 Introduction

This work answers a question raised in [1], namely, identifying the basic ingredients necessary for developing a variational principle all, or some, of whose Euler-Lagrange (E-L) equations are a given system of PDE; the functional to be developed is required to have space-time derivatives of its fundamental fields in 'more-than-linear' combinations. Such a question arose from the purely practical issue of developing a basis for application of Effective Field Theory techniques in Physics (cf. [2, 3, 4]) to the system of nonlinear dislocation dynamics [1] in continuum mechanics and materials science. Despite the success in formulating an appropriate action functional, that effort also exposed a certain flexibility in the adopted scheme whose details remained to be understood. Here, we are able to understand those details and abstract out the essence of the technique. The idea is then demonstrated on a wider setting of important classes of physical systems of nonlinear PDE. We note here that the question of finding a variational principle(s) (some of) whose E-L equations are a given system of PDE is different from the one adopted in the 'Least-Squares Method,' (cf. [5, Ch. 10]), as explained in [6, Sec. 1] and [7]; the E-L equations of the Least Squares functional are not necessarily the PDE system from which the Least-Squares functional is developed. With the null minimizer requirement, minimizers of the Least Squares functional are solutions of the PDE system involved[1]. In the approach adopted herein, a *family* of functionals is developed which satisfy the stated requirement. Mathematically rigorous considerations of the Least-Squares approach rely strongly on convex duality - the present approach relies on elements of convex duality even at a formal level.

    This paper is organized as follows: in the following paragraph of this section we motivate the main idea of this work through an easy computation. In Secs. 2-5 we still use the 'ad-hoc' procedure adopted in [1] to demonstrate our ideas in the context of the incompressible Navier-Stokes (N-S) equations for a homogeneous fluid. Related work can be found in [8, 9] which review earlier work in the physics and mechanics literature, including that of mathematicians C. Doering

---

*Department of Civil & Environmental Engineering, and Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, PA 15213, email: acharyaamit@cmu.edu.

[1]As a related example, it is also instructive to consider the following: for $F : \mathbb{R}^2 \to \mathbb{R}$ let $F(u, \partial_x u) = 0$ be a PDE for $u : (0,1) \to \mathbb{R}$. Let $H : \mathbb{R} \to \mathbb{R}$ be a function with an absolute minimum attained at 0, and consider the functional $S[u] := \int_{(0,1)} dx\, H(F(u, \partial_x u))$. Clearly any solution of the PDE is a minimizer of S. However, the Euler-Lagrange equation of S is $\partial_F H\, \partial_u F - \partial_x(\partial_F H\, \partial_{(\partial_x u)} F) = 0$, which is of course not the same as the original PDE in any sense. I thank Gautam Iyer for a discussion of this case.

and P. Constantin; a sampling of the mathematical literature on the matter can be found in [10, 11] and references provided therein (an important disclaimer here is that I am in no way an expert in the theory or practice of the N-S equations). In Sec. 6, the scheme is generalized to understand the fundamental restrictions, at least at a formal level, required to make it work, exploiting its full level of flexibility; this is demonstrated in the context of a first-order system of conservation laws in divergence form and a system of Hamilton-Jacobi equations. Some closing observations are recorded in Sec. 7. A word on notation: the Einstein summation convention for repeated indices is always invoked; Ω will represent a fixed domain in ambient, 3-d Euclidean space, and $[0, T]$ an interval of time.

Before moving on to the physical models we wish to discuss, we discuss the 'toy' model of the heat equation in one space dimension and time to give the pattern of the typical computations that are involved in each case that follows, only nonlinear and in more space dimensions. Consider the heat equation

$$\partial_t \theta = \partial_x (k \partial_x \theta)$$

and the functional

$$\widehat{S}[\theta, \lambda] := \int_{I \times [0,T]} dt dx \ \lambda(\partial_t \theta - \partial_x(k \partial_x \theta)) + H(\theta)$$

$$= \int_{I \times [0,T]} dt dx \ \theta(-\partial_t \lambda - \partial_x(k \partial_x \lambda)) + H(\theta),$$

for *any* convex function H so that $p = H'(\theta)$ is uniquely solvable for $\theta(p)$, and assuming that λ vanishes on the boundary of the interval $I \subset \mathbb{R}$ (representing the spatial domain). The goal is to propose an 'action' functional whose E-L equation is the original (system of) PDE in question, here the heat equation. Here, we have imposed the PDE (system) with a Lagrange multiplier field (to generate a scalar), exposed linear terms in the basic primal field(s) (here θ) and added a convex term in the basic field, the latter two actions in anticipation of performing a Legendre transform. We note that the Lagrange multiplier field necessarily enters in a 'linear' manner in the functional, even when the PDE is nonlinear (when the PDE contains nonlinear terms - as in the models considered later - all such nonlinear terms are combined additively with the function H to define a function M). Motivated by the structure of $\widehat{S}$, now define

$$M^*(p) := p \, \theta(p) - H(\theta(p))$$
$$p := -(-\partial_t \lambda - \partial_x(k \partial_x \lambda)).$$

With the above definitions, define the functional on a reduced state space

$$S[\lambda] = \int_{I \times [0,T]} dt dx \ - M^*(p)$$

whose first variation is given by

$$\delta S = \int_{I \times [0,T]} dt dx \ - \partial_p M^*(p) \delta p.$$

Noting that $\partial_p M^*(p) = \theta(p)$,

$$\delta S = \int_{I \times [0,T]} dt dx \ \theta(p) \left(-\partial_t \delta \lambda - \partial_x (k \partial_x \delta \lambda)\right)$$

yielding the E-L equation

$$\partial_t(\theta(p)) - \partial_x (k \partial_x(\theta(p))) = 0.$$

With $H(\theta) = \frac{1}{2}\theta^2$, $M^*(p) = \frac{1}{2}p^2$ and $\theta(p) = p$, and the functional $S[\lambda]$ is bounded above. The Euler-Lagrange equation in terms of the function λ is $\partial_t^2\lambda - k^2\partial_x^4\lambda = 0$ which has the wrong sign when viewed as an initial value problem in time t with dispersion relationship $\omega^2 = -k^2m^4$ and eigenmode with wave number m associated with growth factors $e^{\pm|k|m^2 t}$. However, the functional can be maximized or its critical points approached by a gradient ascent in a 'fake time' variable, say s, given by $\partial_s\lambda = \frac{\delta S}{\delta \lambda}$, and it can be expected that the decaying solution (in t) is automatically picked up by such methods. For $k > 0$, it can be checked that initial and boundary conditions on θ can be translated, non-uniquely, to constraints on the field λ at the domain space-time boundaries. One also would seem to have the formal guarantee that any solution in the dual variable λ that respects the boundary and initial constraints obtained from the primal problem must generate the unique solution to the primal problem through the mapping $\theta(x,t) = p(x,t) = \partial_t\lambda(x,t) + k\,\partial_x^2\lambda(x,t)$ (the mapping varies, of course, depending on the choice of the convex function H, but this conclusion remains unaltered).

We note that it is not our intent with the remarks above to suggest that the heat equation be solved in this manner, but only to motivate how the variational principles we develop henceforth may be 'solved' or approximated.

After completing this work, we became aware of the work of Brenier [12]. There are strong connections of the ideas presented in this paper with those in [12] - for instance, the functional (1) with $\hat{\nu} = 0$, $c = 1$, and $G(x) = \frac{1}{2}dx^2, d \to 0_+$ so that $G^*(\xi) = \frac{1}{2d}\xi^2$ formally recovers Brenier's variational principle for the Euler equation (up to accounting for the initial condition). Brenier does rigorous analysis to show existence of generalized solutions to his maximization principle which may be considered as encouraging for the ideas presented herein (the mathematical analysis in [12] is beyond the scope of this paper).

2 A 'dual' action for the incompressible Navier-Stokes equation

For the fields
$$\gamma : \Omega \times [0,T] \to \mathbb{R}$$
$$\lambda : \Omega \times [0,T] \to \mathbb{R}^3$$

consider the functional

$$S_d[\gamma, \lambda] = \int_{\Omega \times [0,T]} dt\, d^3x - \frac{1}{2}p_i \mathbb{K}_{ij}p_j - G^*(\xi), \tag{1}$$

where

$$p_k := -[\hat{\nu}(\partial_j\partial_j\lambda_k + \partial_k\partial_i\lambda_i) - \partial_k\gamma + \partial_t\lambda_k]$$
$$\xi := -\partial_i\lambda_i$$
$$\mathbb{L}_{ij}(\nabla\lambda) := c\,\delta_{ij} + \partial_j\lambda_i + \partial_i\lambda_j; \qquad \mathbb{K}_{ij}(\nabla\lambda) := (\mathbb{L}(\nabla\lambda))_{ij}^{-1}; \qquad \mathbb{K}_{ij} = \mathbb{K}_{ji} \tag{2}$$
$$\omega(\xi) := G'^{-1}(\xi)$$
$$G^*(\xi) := \omega(\xi)\xi - G(\omega(\xi))$$

for $c > 0$ an *arbitrary*, non-dimensional constant, $G : \mathbb{R} \to \mathbb{R}$ an *arbitrary convex* (smooth) function, and $G' : \mathbb{R} \to \mathbb{R}$ refers to the derivative of the function G whose inverse function exists. *We make the assumption here that $\mathbb{L}$ is an invertible matrix for all possible values of its argument* (presumably this can be arranged by taking a sufficiently large value of the constant c). For $\nu > 0$ the shear viscosity and $\rho_0 > 0$ the constant density of the homogeneous, incompressible fluid, $\hat{\nu} := \frac{\nu}{\rho_0}$.

In what follows, we will not always explicitly write the arguments of $\mathbb{L}, \mathbb{K}, \omega$, and G^*.

The first variation of S_d, assuming all variations vanish on the boundary of $\Omega \times [0,T]^2$ is given by

$$\delta S_d = \int_{\Omega \times [0,T]} dt d^3x \ -p_i \mathbb{K}_{ik}[-(\hat{\nu}(\partial_j \partial_j \delta \lambda_k + \partial_k \partial_i \delta \lambda_i) - \partial_k \delta \gamma + \partial_t \delta \lambda_k)]$$
$$-G^{*\prime}[-\partial_i \delta \lambda_i] \tag{3}$$
$$-p_i p_k \, \delta \mathbb{K}_{ik},$$

and noting that

$$\delta \mathbb{K}_{ik} = -\mathbb{K}_{ij} \delta \mathbb{L}_{jm} \mathbb{K}_{mk} = -\mathbb{K}_{ij} \mathbb{K}_{mk}[\partial_j \delta \lambda_m + \partial_m \delta \lambda_j]$$

along with the definition

$$v_k(p, \nabla \lambda) := \mathbb{K}_{ik}(\nabla \lambda) p_i \tag{4}$$

we have

$$\delta S_d = \int_{\Omega \times [0,T]} dt d^3x \ \hat{\nu}(\delta \lambda_k \partial_j \partial_j v_k + \delta \lambda_i \partial_i \partial_j v_j)$$
$$+ (\partial_k v_k) \delta \gamma - (\partial_t v_k) \delta \lambda_k$$
$$- \frac{1}{2}[\partial_j(v_j v_k) + \partial_m(v_k v_m)] \delta \lambda_k$$
$$- \partial_i(G^{*\prime}) \delta \lambda_i.$$

Noting that

$$G^*(\xi) = \omega(\xi)\xi - G(\omega(\xi)) \ \text{ and } \ G'(\omega(\xi)) = \xi \Longrightarrow G^{*\prime}(\xi) = \omega(\xi)$$

and defining

$$\mathcal{P} := \rho_0 \, \omega \tag{5}$$

the following Euler-Lagrange equations are obtained:

$$\delta \lambda_k : \qquad -\partial_t v_k - \partial_j(v_k v_j) + \partial_j \left(\hat{\nu}(\partial_j v_k + \partial_k v_j) - \frac{\mathcal{P}}{\rho_0} \delta_{kj} \right) = 0 \tag{6}$$
$$\delta \gamma : \qquad \partial_k v_k = 0.$$

These are the Navier-Stokes system for an incompressible, homogeneous fluid if the fields $v, \mathcal{P}$ (that are defined in terms of the fields λ and γ) are interpreted as the *velocity* and the *pressure* fields, respectively.

3 The primal action for (1) and its reduced state space

We motivate how the functional (1) was arrived at. With all definitions and notation of the previous section enforced and in terms of the fields

$$v : \Omega \times [0,T] \to \mathbb{R}^3$$
$$\gamma : \Omega \times [0,T] \to \mathbb{R}$$
$$\lambda : \Omega \times [0,T] \to \mathbb{R}^3$$
$$\omega : \Omega \times [0,T] \to \mathbb{R}$$

---

[2]Here, we are interested in interior field equations; natural 'boundary' conditions can be inferred by not assuming the variations to vanish on the boundary of the space-time domain and appropriately utilizing applied initial and spatial boundary conditions of the primal problem.

consider the functional

$$\widehat{S}_d[v, \gamma, \lambda, \omega] = \int_{\Omega \times [0,T]} dt d^3x \; c\frac{1}{2}v_i v_i + G(\omega)$$
$$+ \gamma(\partial_i v_i)$$
$$+ \lambda_i(\partial_j(-\omega\delta_{ij} + \hat{\nu}(\partial_j v_i + \partial_i v_j)) - \partial_t v_i - \partial_j(v_i v_j)),$$

where the incompressible Navier-Stokes equations have been enforced via Lagrange multipliers.

Each term in the action density has physical dimensions of $\frac{Length^2}{Time^2}$; multiplying through by ρ_0 gives units of energy per unit volume.

The main idea is to invoke a Legendre transform based change of variables and then consider the variational principle in a *reduced state space*. Assuming for the moment that the Lagrange multiplier fields vanish on the boundary, we have

$$\widehat{S}_d[v, \gamma, \lambda, \omega] = \int_{\Omega \times [0,T]} dt d^3x \; c\frac{1}{2}v_i v_i + v_i v_j \partial_j \lambda_i + G(\omega)$$
$$+ (\hat{\nu}(\partial_j \partial_j \lambda_i + \partial_i \partial_k \lambda_k) - \partial_i \gamma + \partial_t \lambda_i)v_i$$
$$+ \omega \partial_i \lambda_i.$$

Working with the definitions (2) and (4), we affect a reduction in the state space of $\widehat{S}$ to define

$$S_d[\gamma, \lambda] = \int_{\Omega \times [0,T]} dt d^3x \; \frac{1}{2}v_i(p, \nabla\lambda)\mathbb{L}_{ij}(\nabla\lambda)v_j(p, \nabla\lambda) - p_i v_i(p, \nabla\lambda) + G(\omega(\xi)) - \omega(\xi)\xi. \qquad (7)$$

Using the definitions (2) and (4) once again and noting that

$$v_i p_i - \frac{1}{2}v_i \mathbb{L}_{ij} v_j = \frac{1}{2}p_i \mathbb{K}_{ik} p_k, \qquad (8)$$

we note that (7) is the functional (1).

4 A 'mixed' action for the incompressible Navier-Stokes equation

For the fields

$$A : \Omega \times [0,T] \to \mathbb{R}^{3\times3}_{sym}$$
$$\gamma : \Omega \times [0,T] \to \mathbb{R}$$
$$\lambda : \Omega \times [0,T] \to \mathbb{R}^3$$
$$\omega : \Omega \times [0,T] \to \mathbb{R}$$

consider the functional

$$S_m[A, \gamma, \lambda, \omega] = \int_{\Omega \times [0,T]} dt d^3x \; -\frac{1}{2}p_i \mathbb{K}_{ij} p_j \mp R^*(\tau) + \omega \partial_i \lambda_i, \qquad (9)$$

where

$$p_k := -[-\partial_j A_{kj} - \partial_k \gamma + \partial_t \lambda_k]$$
$$\tau_{kl} := \mp[-A_{kl} - \hat{\nu}(\partial_l \lambda_k + \partial_k \lambda_l)]$$
$$\mathbb{L}_{ij}(\nabla\lambda) := c\,\delta_{ij} + \partial_j \lambda_i + \partial_i \lambda_j; \qquad \mathbb{K}_{ij}(\nabla\lambda) := (\mathbb{L}(\nabla\lambda))^{-1}_{ij}; \qquad \mathbb{K}_{ij} = \mathbb{K}_{ji} \qquad (10)$$
$$D_{ij}(\tau) := \left(R'^{-1}\right)_{ij}(\tau)$$
$$R^*(\tau) := D_{ij}(\tau)\tau_{ij} - R(D(\tau))$$

5

for $R : \mathbb{R}^{3\times 3}_{sym} \to \mathbb{R}$ an arbitrary *convex* function on the space of symmetric tensors, and $R'_{ij} : \mathbb{R}^{3\times 3}_{sym} \to \mathbb{R}^{3\times 3}_{sym}$ refers to the function $\partial_{D_{ij}} R$ and R'^{-1}_{ij} to its inverse function.

The matrix field $\mathbb{L}$ and the constant $\hat{\nu}$ are defined exactly as in Sec. 2.

The first variation of S_m, assuming all variations vanish on the boundary of $\Omega \times [0, T]$ (with the same understanding for what needs to be done to include natural boundary and initial conditions) is given by

$$\delta S_m = \int_{\Omega \times [0,T]} dt d^3 x \; -p_i \mathbb{K}_{ik}(\nabla \lambda)[-(-\partial_j \delta A_{kj} - \partial_k \delta \gamma + \partial_t \delta \lambda_k)]$$
$$\mp R^{*'}_{kl}[\mp(-\delta A_{kl} - 2\hat{\nu}\partial_l \delta \lambda_k)] \tag{11}$$
$$+ \delta \omega \partial_i \lambda_i - (\partial_i \omega)\delta \lambda_i - \frac{1}{2}p_i p_k \delta \mathbb{K}_{ik},$$

and noting that

$$\delta \mathbb{K}_{ik} = -\mathbb{K}_{ij}\delta \mathbb{L}_{jm}\mathbb{K}_{mk}$$

we have

$$\delta S_m = \int_{\Omega \times [0,T]} dt d^3 x \; \partial_j(p_i \mathbb{K}_{ik}(\nabla \lambda))\delta A_{kj} + \partial_k(p_i \mathbb{K}_{ik}(\nabla \lambda))\delta \gamma - \partial_t(p_i \mathbb{K}_{ik}(\nabla \lambda))\delta \lambda_k$$
$$- R^*_{kl}\delta A_{kl} + \partial_l(2\hat{\nu}R^{*'}_{kl})\delta \lambda_k + (\partial_i \lambda_i)\delta \omega$$
$$- \frac{1}{2}[\delta \lambda_m \partial_j(\mathbb{K}_{ij}(\nabla \lambda)p_i \mathbb{K}_{mk}(\nabla \lambda)p_k) + \delta \lambda_j \partial_m(\mathbb{K}_{ij}(\nabla \lambda)p_i \mathbb{K}_{mk}(\nabla \lambda)p_k)],$$

leading to the Euler-Lagrange equations

$$\delta A_{ik}: \quad \frac{1}{2}[\partial_j(p_i \mathbb{K}_{ik}(\nabla \lambda)) + \partial_k(p_i \mathbb{K}_{ij}(\nabla \lambda))] - R^{*'}_{kj}(\tau) = 0$$
$$\delta \gamma: \quad \partial_k(p_i \mathbb{K}_{ik}(\nabla \lambda)) = 0$$
$$\delta \lambda_k: \quad -\partial_t(p_i \mathbb{K}_{ik}(\nabla \lambda)) + \partial_l(2\hat{\nu}R^{*'}_{kl}(\tau)) - \partial_k \omega \tag{12}$$
$$- \frac{1}{2}[\partial_j(\mathbb{K}_{ij}(\nabla \lambda)p_i \mathbb{K}_{km}(\nabla \lambda)p_m) + \partial_m(\mathbb{K}_{ik}(\nabla \lambda)p_i \mathbb{K}_{mr}(\nabla \lambda)p_r)] = 0$$
$$\delta \omega: \quad \partial_i \lambda_i = 0.$$

Defining a velocity and a pressure field as

$$v_k := \mathbb{K}_{ik}p_i; \qquad \mathcal{P} := \rho_0 \omega \tag{13}$$

we note that the first three equations of (12) imply the equations

$$\partial_k v_k = 0$$
$$-\partial_t v_k - \partial_j(v_k v_j) + \partial_j\left[\frac{\nu}{\rho_0}(\partial_j v_k + \partial_k v_j) - \frac{\mathcal{P}}{\rho_0}\delta_{kj}\right] = 0, \tag{14}$$

which is the Navier-Stokes system for a homogeneous, incompressible fluid.

5 The primal action for (9) and its reduced state space

We motivate how the functional (9) was arrived at. With all definitions and notation of previous sections enforced and in terms of the fields

$$v : \Omega \times [0, T] \to \mathbb{R}^3$$
$$D : \Omega \times [0, T] \to \mathbb{R}^{3\times 3}_{sym}$$

consider the functional

$$\widehat{S}_m[v, D, \omega, A, \gamma, \lambda] = \int_{\Omega \times [0,T]} dt d^3x \, c\frac{1}{2}v_i v_i \pm R(D)$$
$$+ A_{ij}(\partial_j v_i - D_{ij})$$
$$+ \gamma(\partial_i v_i)$$
$$+ \lambda_i(\partial_j(-\omega\delta_{ij} + 2\hat{\nu}D_{ij}) - \partial_t v_i - \partial_j(v_i v_j)),$$

where the incompressible Navier-Stokes equations have been enforced via Lagrange multipliers.

As before, we invoke a Legendre transform based change of variables and then consider the variational principle in a reduced state space. Assuming for the moment that the Lagrange multiplier fields vanish on the boundary, we have

$$\widehat{S}_m[v, D, \omega, A, \gamma, \lambda] = \int_{\Omega \times [0,T]} dt d^3x \, c\frac{1}{2}v_i v_i \pm R(D)$$
$$+ (-\partial_j A_{ij} - \partial_i \gamma + \partial_t \lambda_i)v_i$$
$$+ (-A_{kl} - \hat{\nu}(\partial_l \lambda_k + \partial_k \lambda_l))D_{kl}$$
$$+ \omega\partial_i \lambda_i + v_i v_j \partial_j \lambda_i.$$

Working with the definitions (10) and (13), we affect a reduction in the state space of $\widehat{S}_m$ to define

$$S_m[A, \gamma, \lambda, \omega] = \int_{\Omega \times [0,T]} dt d^3x \, \frac{1}{2}v_i(p, \nabla\lambda)\mathbb{L}_{ij}(\nabla\lambda)v_j(p, \nabla\lambda) - p_i v_i(p, \nabla\lambda)$$
$$\pm R(D(\tau)) \mp \tau_{kl}D_{kl}(\tau) + \omega\partial_i \lambda_i. \tag{15}$$

Using the definitions (10) and (13) once again and noting (8), we note that (15) is the functional (9).

6   Generalizations

In the previous sections and in [1] the kinetic energy was chosen as the added potential associated with the velocity field in the primal actions. It is however clear that apart from a solvability condition, this potential should be amenable to a (more or less) arbitrary choice. In this section we demonstrate this feature of the proposed scheme through a discussion of important classes of PDE.

In the following subsections, we will repeatedly make use of an *assumption* and its consequence, which we write out in detail before proceeding. It is essentially related to a Legendre transform in the presence of a parameter. For $M : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $H : \mathbb{R}^n \to \mathbb{R}$, $L \in \mathbb{R}^m$, and $F : \mathbb{R}^n \to \mathbb{R}^m$ satisfying

$$M(U, L) = H(U) - L \cdot F(U), \tag{16}$$

we assume that for given P and L, there exists a unique function $U(P, L)$ which satisfies the relation

$$P = \partial_U M(U(P, L), L). \tag{17}$$

for all likely values of L and P to be encountered, i.e., the algebraic system of equations $P = \partial_U M(U, L)$ is uniquely solvable for U in terms of P, L. If M were to be convex in U for all L, then such a condition would certainly hold. In the following, the function F will be specified from the

physical PDE system to be solved and the function H will be free to choose, so this condition is essentially a constraint on the choice of the class of functions to which H belongs.

Assuming the above, consider

$$M^*(P, L) := U(P, L) \cdot P - M(U(P, L), L)$$
$$\implies \partial_L M^*(P, L) = \partial_L U_i(P, L) P_i - \partial_{U_i} M(U(P, L), L) \partial_L U_i(P, L) - \partial_L M(U(P, L), L) \tag{18}$$
$$= -\partial_L M(U(P, L), L),$$

due to (17). Given the form of (16), we have

$$\partial_L M^*(P, L) = F(U(P, L)). \tag{19}$$

We also have that

$$\partial_P M^*(P, L) = \partial_P U_i(P, L) P_i + U(P, L) - \partial_{U_i} M(U(P, L), L) \partial_P U_i(P, L) = U(P, L) \tag{20}$$

by (17).

6.1 Dual variational principles for first-order systems of conservation laws in divergence form

The PDE system of interest is of the form

$$\partial_t u_I + \partial_i f_{Ii}(u) = 0, \tag{21}$$

where $I = 1, \ldots, n$, and $i = 1, \ldots, d$, d being the number of space dimensions.

In terms of the field

$$\lambda : \Omega \times [0, T] \to \mathbb{R}^n$$

consider the functional

$$S_{cl}[p, \lambda] = \int_{\Omega \times [0,T]} dt d^3 x - M^*(p, \nabla \lambda), \tag{22}$$

where $(\nabla \lambda)_{Ij} = \partial_j \lambda_I$, along with the identifications $u = U$, $p = P$, $\nabla \lambda = L$, $m = nd$, $F = f$ in (16)-(17) and the definitions

$$p_I := \partial_t \lambda_I$$
$$M(u(p, \nabla \lambda), \nabla \lambda) := H(u(p, \nabla \lambda)) - \nabla \lambda \cdot f(u(p, \nabla \lambda)) \tag{23}$$
$$M^*(p, \nabla \lambda) := u(p, \nabla \lambda) \cdot p - M(u(p, \nabla \lambda), \nabla \lambda)$$

for *any* choice of the function H that allows (17) to hold. Then the first variation is given by

$$\delta S_{cl} = \int_{\Omega \times [0,T]} dt d^3 x - (\partial_p M^* \cdot \partial_t \delta \lambda + \partial_{\nabla \lambda} M^* \cdot \nabla \delta \lambda)$$

which leads to the E-L equations

$$\partial_t u_I(p, \nabla \lambda) + \partial_i f_{Ii}(u(p, \nabla \lambda)) = 0$$

by (19)-(20). Thus, every solution of the E-L equations of (22) defines a solution of (21) through the definition of the function $U(P, L)$ related to (17).

We also note that for given f, each member of the entire class of functions M^*, defined through a choice of an admissible state function H, satisfies the conservation law (21) in the sense

$$\partial_t (\partial_p M^*)_I + \partial_i (\partial_{\nabla \lambda} M^*)_{Ii} = 0. \tag{24}$$

8

6.2 Dual variational principle for a second order system of Hamilton-Jacobi equations

Consider the system of Hamilton-Jacobi equations

$$\partial_t u_I = f_I(u, B, C)$$
$$\partial_i u_I = B_{Ii} \tag{25}$$
$$\partial_i \partial_j u_I = C_{Iji}$$

where f_I is a smooth function of its arguments.

In terms of the fields

$$\lambda : \Omega \times [0, T] \to \mathbb{R}^n$$
$$\gamma : \Omega \times [0, T] \to \mathbb{R}^{n \times d}$$
$$\rho : \Omega \times [0, T] \to \mathbb{R}^{n \times d \times d}$$

and the definitions

$$P := (\partial_t \lambda + \nabla \cdot \gamma - \nabla^2 : \rho, \gamma, \rho)$$
$$L := \lambda \tag{26}$$
$$F := f,$$

where $(\nabla \cdot \gamma)_I = \partial_i \gamma_{Ii}$ and $(\nabla^2 : \rho)_I = \partial_j \partial_i \rho_{Iij}$, we consider functions $H : \mathbb{R}^n \times \mathbb{R}^{n \times d} \times \mathbb{R}^{n \times d \times d} \to \mathbb{R}$ such that, for the generic element in its domain referred to as

$$U := (u, B, C),$$

the function M in (16) is defined with the solvability property (17), and in terms of it, the function M^* in (18).

Consider the functional

$$S_{HJ}[\lambda, \gamma, \rho] = \int_{\Omega \times [0,T]} dt d^3 x - M^*(P, L), \tag{27}$$

whose E-L equations are (assuming (λ, γ, ρ) have compact support on Ω)

$$-\partial_t u_I(P, L) + f_I(U(P, L)) = 0$$
$$-\partial_i u_I(P, L) + B_{Ii}(P, L) = 0$$
$$-\partial_i \partial_j u_I(P, L) + C_{Iji}(P, L) = 0.$$

We note that the static system $\partial_i f_{Ii}(B) = 0$, $B_{Ii} = \partial_i u_I$ can be dealt with as an Hamilton-Jacobi system as well as by taking account of its conservation structure by the proposed technique, with $L = \nabla \lambda$ in the latter case.

7 Concluding remarks

The proposed scheme for generating variational principles for nonlinear PDE systems may be abstracted as follows: We first pose the given system of PDE as a *first-order* system (introducing extra fields representing (higher-order) space and time derivatives of the fields of the given system); as before let us denote this collection of primal fields by U. 'Multiplying' the primal equations by dual Lagrange multiplier fields, the collection denoted by D, adding a function $H(U)$, solely in the variables U (the purpose of which, and associated requirements, will be clear shortly), and

integrating by parts over the space-time domain, we form a 'mixed' functional in the primal and dual fields given by

$$\widehat{S}_H[U, D] = \int_{[0,T] \times \Omega} dt d^3 x \; \mathcal{L}_H(\mathcal{D}, U)$$

where $\mathcal{D}$ is a collection of local objects in D and at most its first order derivatives. We then require that the family of functions H be such that it allows the definition of a function $U_H(\mathcal{D})$ such that

$$\frac{\partial \mathcal{L}_H}{\partial U}(\mathcal{D}, U_H(\mathcal{D})) = 0$$

so that the *dual* functional, defined solely on the space of the dual fields D, given by

$$S_H[D] = \int_{[0,T] \times \Omega} dt d^3 x \; \mathcal{L}_H(\mathcal{D}, U_H(\mathcal{D}))$$

has the first variation

$$\delta S_H = \int_{[0,T] \times \Omega} dt d^3 x \; \frac{\partial \mathcal{L}_H}{\partial \mathcal{D}} \delta \mathcal{D}.$$

By the process of formation of the functional $\widehat{S}_H$, it can then be seen that the (formal) E-L equations arising from δS_H have to be the original first-order primal system, with U substituted by $U_H(\mathcal{D})$, regardless of the H employed.

Thus, the proposed scheme may be summarized as follows: we wish to pursue the following (local-global) critical point problem

$$\underset{D}{\text{extremize}} \int_{[0,T] \times \Omega} dt d^3 x \; \underset{U}{\text{extremize}} \; \mathcal{L}_H(\mathcal{D}(t, x), U),$$

where the pointwise extremization of $\mathcal{L}_H$ over U, for fixed $\mathcal{D}$, is made possible by the choice of H.

Furthermore, assume the Lagrangian $\mathcal{L}_H$ can be expressed in the form

$$\mathcal{L}_H(\mathcal{D}, U) := -P(\mathcal{D}) \cdot U + f(U, D) + H(U)$$

for some function P defined by the structure of the primal first-order system ((linear terms in) first derivatives of U after multiplication by the dual fields and integration by parts always produce such terms), and for some function f which, when non-zero, does not contain any linear dependence in U. Our scheme requires the existence of a function U_H defined from 'solving $\frac{\partial \mathcal{L}}{\partial U}(\mathcal{D}, U) = 0$ for U,' i.e. $\exists\, U_H(P(\mathcal{D}), \mathcal{D})$ s.t. the equation

$$-P(\mathcal{D}) + \frac{\partial f}{\partial U}(U_H(P(\mathcal{D}), \mathcal{D}), \mathcal{D}) + \frac{\partial H}{\partial U}(U_H(P(\mathcal{D}), \mathcal{D})) = 0$$

is satisfied. This requirement may be understood as follows: define

$$f(U, \mathcal{D}) + H(U) =: M(U, \mathcal{D})$$

and assume that it is possible, through the choice of H, to make the function $\frac{\partial M}{\partial U}(U, \mathcal{D})$ *monotone* in U so that a function $U_H(P, \mathcal{D})$ can be defined that satisfies

$$\frac{\partial M}{\partial U}(U_H(P, \mathcal{D}), \mathcal{D}) = P, \quad \forall P.$$

Then the Lagrangian is

$$\mathcal{L}(\mathcal{D}, U_H(P(\mathcal{D}), \mathcal{D})) = -P(\mathcal{D}) \cdot U_H(P(\mathcal{D}), \mathcal{D}) + M(U_H(P(\mathcal{D}), \mathcal{D}), \mathcal{D}) =: -M^*(P(\mathcal{D}), \mathcal{D})$$

where $M^*(P, \mathcal{D})$ is the Legendre transform of the function M w.r.t U, with $\mathcal{D}$ considered as a parameter.

Thus, our scheme may also be interpreted as designing a concrete realization of abstract saddle point problems in optimization theory [13], where we exploit the fact that, in the context of 'solving' PDE viewed as constraints implemented by Lagrange multipliers to generate an unconstrained problem, there is a good deal of freedom in choosing an objective function(al) to be minimized. We exploit this freedom in choosing the function H to develop dual variational principles corresponding to general systems of PDE.

We conclude with a few observations and directions for future work:

1. A detailed study of how boundary and initial conditions of the primal problem can be transferred to the same for the dual problem is warranted, most importantly for practical purposes of generating numerical approximations to the proposed formal mathematical scheme.

   As a first example, consider the dual formulation of the heat equation discussed in Sec. 1. It is already clear that such a problem in space-time cannot be viewed as an initial value problem because of the lack of continuous dependence w.r.t initial data, as discussed. Thus, one has to consider the dual problem as a boundary-value problem in space-time. An important question to resolve then is whether doing so allows enough freedom for predicting the correct evolving, primal temperature profile. For this, it seems reasonable that the dual E-L equation $(\partial_t^2 \lambda - k^2 \partial_x^4 \lambda = 0)$, admits two side conditions in the time (t) direction, and four conditions in the space (x) direction. For a unique evolution on the primal side, one has to apply the boundary conditions of the primal problem, say pure Dirichlet on temperature θ, and an initial condition on θ as well, which translate to constraints on the fields of the dual problem through the mapping, $\theta(x,t) = \partial_t \lambda(x,t) + k \partial_x^2 \lambda(x,t)$. Thus, for a solution of the dual problem in the space-time domain $(x_l, x_r) \times [0, T]$, specifying $\lambda(x, T)$ (arbitrarily) and $\theta(x, 0)$ along with $\lambda(x_l, t), \lambda(x_r, t)$ (arbitrarily), and $\theta(x_l, t), \theta(x_r, t)$ would seem to be a possibility, which should still allow enough freedom in the development of $\theta(x, T)$ through the mapping $\theta(x, T) = \partial_t \lambda(x, T) + k \partial_x^2 \lambda(x, T)$. Such questions are the focus of ongoing work.

2. Setting $\hat{\nu} = 0$ in the functionals (1) and (9) yields stationary principles for the incompressible Euler equations.

3. The mixed variational principle for the incompressible N-S equations in Secs. 4-5 affords the addition of added viscosity through the function R in the solution of the dual problem - it is an interesting question whether this feature of the problem can help in the analysis and solution of problems of turbulence.

4. The pointwise invertibility of the matrix field $\mathbb{L}$ in Secs. 2-5 appears to be a key issue in the formalism. It would be interesting to understand the effect of the condition $\det(\mathbb{L}) = 0$ on the proposed scheme and whether there can be a relation between the Reynolds number and the value of c for optimal performance of the scheme.

   In the general setting, this condition translates to the validity of the assumption (17).

5. Constrained by appropriate boundary and initial conditions, when the PDE system has unique solutions, it is clear that *any* choice of the function H within the admissible class, coupled with the mapping from the dual fields to the primal fields, must lead to the unique solution. In the absence of uniqueness, it is an interesting question whether specific choices of H (along with the dual-to-primal mapping) act as a selection mechanism for picking up particular solutions. For instance, an appropriate quadratic choice of H for the 'simple' equation $\partial_x f(B) = 0, B =$

$\partial_x u$ where f is nonconvex results in an essentially semilinear second-order dual problem that is definitely simpler than the quasilinear primal problem. The structure of the dual problem in this case does not make the expectation of a smooth solution an absurd one. If this is indeed borne out in reality, then the solution to the primal problem that is defined through the dual-to-primal mapping may also be expected to be smooth.

6. In the context of first-order systems of conservation laws, what connection, if any, might exist between the large class of functions M^* and entropies [14] of conservation laws is an important question to resolve.

7. There does not exist well-established computational approximation schemes (and for that matter, theory [15]) for *systems* of H-J equations. This work provides a variational structure for such systems which naturally lends itself to, say a finite element, discretization. Whether such an idea has any practical merit would be interesting to explore. If so, this can be very useful for certain systems related to the mechanics of fracture and plastic deformation of solids, see [16, 17, 18].

8. For static conservation laws of the form $\partial_i f_{Ii}(\nabla u) = 0, (\nabla u)_{Ij} := \partial_j u_I$, the scheme produces variational principles even when there does not exist a potential $\psi(\nabla u)$ such that $f_{Ii} = \partial_{(\nabla u)_{Ii}} \psi$.

9. The considerations herein show that even a given variational principle can be associated with a vastly different dual variational principle through the choice of the function H, by associating the former's E-L equations with the latter following the proposed scheme. This seems to open up fascinating points of convergence between apparently different classes of physical models described in terms of PDE and/or variational principles. A particularly intriguing question is whether the present considerations have any connection to the correspondences like AdS-CFT or AdS-CMT [19, 20] in the string theory-high energy-condensed matter-gravitational physics nexus.

10. The proposed duality scheme seems to suggest that for a given PDE system, elements of the class of admissible potentials H form symmetry operations for the system. The algebraic structure of this class of potentials, starting from whether they form a group, is of independent interest, as well as whether knowledge of such structure can help in the understanding of solutions to the PDE system.

Acknowledgments

References

[1] A. Acharya, "An action for nonlinear dislocation dynamics," *Journal of the Mechanics and Physics of Solids*, vol. 161, p. 104811, 2022.

[2] H. Kleinert, *Gauge fields in condensed matter. Vol I: Superflow and Vortex lines.* Singapore Teaneck, N.J: World Scientific, 1989.

[3] A. J. Beekman, J. Nissinen, K. Wu, K. Liu, R.-J. Slager, Z. Nussinov, V. Cvetkovic, and J. Zaanen, "Dual gauge field theory of quantum liquid crystals in two dimensions," *Physics Reports*, vol. 683, pp. 1–110, 2017.

[4] A. J. Beekman, J. Nissinen, K. Wu, and J. Zaanen, "Dual gauge field theory of quantum liquid crystals in three dimensions," *Physical Review B*, vol. 96, no. 16, p. 165115, 2017.

[5] S. Armstrong, T. Kuusi, and J.-C. Mourrat, *Quantitative stochastic homogenization and large-scale regularity.* Springer, 2019, vol. 352, Grundlehren der mathematischen Wissenschaften.

[6] R. L. Seliger and G. B. Whitham, "Variational principles in continuum mechanics," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 305, no. 1480, pp. 1–25, 1968.

[7] A. Visintin, "Minimization vs. null-minimization: a note about the Fitzpatrick theory," https://www.mathematik.tu-darmstadt.de/media/isimm/ISIMM-Forum_Visintin_1409.pdf, 2014.

[8] R. Kerswell, "Variational principle for the Navier-Stokes equations," *Physical Review E*, vol. 59, no. 5, p. 5482, 1999.

[9] G. Liu, "Dual variational principles for 3-d Navier-Stokes equations," in *New Trends in Fluid Mechanics Research.* Springer, 2007, pp. 734–735.

[10] N. Ghoussoub and A. Moameni, "Anti-symmetric Hamiltonians (II): Variational resolutions for Navier-Stokes and other nonlinear evolutions," in *Annales de l'IHP Analyse non linéaire*, vol. 26, no. 1, 2009, pp. 223–255.

[11] M. Ortiz, B. Schmidt, and U. Stefanelli, "A variational approach to Navier–Stokes," *Nonlinearity*, vol. 31, no. 12, p. 5664, 2018.

[12] Y. Brenier, "The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem," *Communications in Mathematical Physics*, vol. 364, no. 2, pp. 579–605, 2018.

[13] R. T. Rockafellar, *Conjugate duality and optimization.* SIAM, 1974.

[14] C. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics.* Springer, 2016, vol. 325, Grundlehren der mathematischen Wissenschaften.

[15] R. Hynd, "The Hamilton-Jacobi equation, then and now," *Notices of the American Mathematical Society*, vol. 68, no. 9, pp. 1457–1467, 2021.

[16] L. Morin and A. Acharya, "Analysis of a model of field crack mechanics for brittle materials," *Computer Methods in Applied Mechanics and Engineering*, vol. 386, p. 114061, 2021.

[17] X. Zhang, A. Acharya, N. J. Walkington, and J. Bielak, "A single theory for some quasi-static, supersonic, atomic, and tectonic scale applications of dislocations," *Journal of the Mechanics and Physics of Solids*, vol. 84, pp. 145–195, 2015.

[18] R. Arora and A. Acharya, "A unification of finite deformation J_2 von-Mises plasticity and quantitative dislocation mechanics," *Journal of the Mechanics and Physics of Solids*, vol. 143, p. 104050, 2020.

[19] J. Zaanen, Y. Liu, Y.-W. Sun, and K. Schalm, *Holographic duality in condensed matter physics*. Cambridge University Press, 2015.

[20] S. A. Hartnoll, A. Lucas, and S. Sachdev, *Holographic quantum matter*. MIT press, 2018.