This paper presents a reasonably complete duality theory and a nonlinear dual transformation method for solving the fully nonlinear, non-convex parametric variational problem
\[
\inf \{ W(\Lambda u - \mu) - F(u) \},
\]
and associated nonlinear boundary value problems, where \( \Lambda \) is a nonlinear operator, \( W \) is either convex or concave functional of \( p = \Lambda u \), and \( \mu \) is a given parameter. Detailed mathematical proofs are provided for the complementary extremum principles proposed recently in finite deformation theory. A method for obtaining truly dual variational principles (without a dual gap and involving the dual variable \( p^* \) of \( \Lambda u \) only) in \( n \)-dimensional problems is proposed. It is proved that for convex \( W(p) \), the critical point of the associated Lagrangian
\[
L_\mu(u, p^*)
\]
is a saddle point if and only if the so-called complementary gap function is positive. In this case, the system has only one dual problem. However, if this gap function is negative, the critical point of the Lagrangian is a so-called super-critical point, which is equivalent to the Auchmuty’s anomalous critical point in geometrically linear systems. We discover that, in this case, the system may have more than one primal-dual set of problems. The critical point of the Lagrangian either minimizes or maximizes both primal and dual problems. An interesting triality theorem in non-convex systems is proved, which contains a minimax complementary principle and a pair of minimum and maximum complementary principles. Applications in finite deformation theory are illustrated. An open problem left by Hellinger and Reissner is solved completely and a pure complementary energy principle is constructed. It is proved that the dual Euler–Lagrange equation is an algebraic equation, and hence, a general analytic solution for non-convex variational-boundary value problems is obtained. The connection between nonlinear differential equations and algebraic geometry is revealed.

**Keywords:** duality theory; triality; complementary energy; nonlinear variational problem; non-convex optimization; fully nonlinear system; finite deformation theory; partial differential equations; analytic solution; phase transitions; variational inequality.

1. Introduction

We are interested in the non-convex parametric variational problem
\[
P_\mu(u) = J(u, \Lambda u - \mu) \to \min \forall u \in U_\mu,
\]
where \( \Lambda \) is a nonlinear differential operator with \( \Lambda(u) \) written simply as \( \Lambda u \) and \( \mu \) is a given distributed parameter. This problem appears in many physical systems such as hysteresis and phase transitions, spinodal decomposition, non-convex optimal design and control, nonlinear bifurcation and stability analysis of finite deformation theory, nonlinear
elasticity with residual strain, post-buckling of large deformed structures and many others. Very often, the total potential $P_\mu$ can be split into two parts, so that $P_\mu(u) = W(Au - \mu) - F(u)$, where $W$ is the internal (or stored) energy, while $F$ is the external energy. For example, in the equilibrium problem of Ericksen’s bar subjected to an axial extension load, the total potential $P_\mu$ is a double-well energy

$$P_\mu(u) = W(Au - \mu) - F(u) = \int_0^1 \left[ \frac{1}{2} E(x)[\frac{1}{2} u_x^2 + c(x) u_x - \mu] \right] dx - \int_0^1 f u dx, \quad (1)$$

where $E > 0$ and $c(x)$ and $f(x)$ are given functions, while, in this problem $A$ is a quadratic operator. It is surprising that in the post-buckling analysis of nonlinear extended beams subjected to an axial compressive load $\mu$, the total potential energy is almost the same as in (1) (Gao 1996a,b, 1997). For a given lateral load $f(x)$, $P_\mu(u)$ may have two minimizers, corresponding to two possible buckling states, and one local maximizer, corresponding to an unstable state.

It seems that the non-convex variational problem with double-well energy was first studied by van der Waals in 1893 for a compressible fluid whose free energy at constant temperature depends not only on the density but also on the density gradient (Rowlinson 1979). Direct approaches and relaxation methods for solving non-convex variational problems have been discussed extensively during the last thirty years; see, for example, (Ball 1977; Ball & James 1987; Carr et al. 1984; Chang 1981; Dacorogna 1989; Dem’yanov et al. 1996; Ericksen 1986; Fonseca 1988; Fried & Gurtin 1996; James & Kinderlehrer 1989; Kohn 1991; Kohn & Strang 1987; Lurie & Cherkaev 1988; Milton 1990; Pipkin 1993; Steigmann & Ogden 1997; Temam & Strang 1980; Wu 1994). But even for very simple one-dimensional problems like (1), to obtain all analytic solutions is usually very difficult with traditional methods, and therefore, some numerical discretization approaches have been examined in Rogers & Truskinovsky (1997). Relaxation methods, however, can be used mainly for obtaining global minimizers of the relaxed problems. In phase transitions or post-bifurcation analysis, local maximizers usually play an important role in understanding the physical mechanism of systems. As was indicated in Kohn (1991), the relaxation method for solving non-convex variational problems with three or more phases (potential wells) is fundamentally more difficult.

In this paper we take a different approach. While we see the intrinsic difficulties in traditional direct approaches and relaxation methods, we will study the duality theory in general non-convex parametric variational problems. Duality theory has always played a crucial role in mathematics and science. For geometrically linear (that is, infinitesimal deformation) systems, where $A$ is a linear operator, duality theory has been well studied for both convex problems (see, for example, Ekeland & Temam 1976; Fenchel 1949; Rockafellar 1967, 1974; Tonti 1972; Walk 1989; Wright 1997) and non-convex problems; see, for example, (Auchmuty 1983, 1989; Toland 1978, 1979). Applications in continuum mechanics and variational inequalities have been given in several well-known books; see, for example, (Oden 1986; Oden & Reddy 1983; Panagiotopoulos 1985, 1993). If the system has a convex stored energy $W$, the symmetry between the primal and dual energy principles is amazingly beautiful; see, for example, (Marsden & Ratiu 1995; Sewell 1987; Strang 1986; Tabarrok & Rimrott 1994). The dual functional $P^*(p^*) = J^*(A^* p^*, p^*)$ in these classical principles is obtained by the well-known Fenchel–Rockafellar duality theory, where $p^*$ is the dual variable of $p = Au$ and $A^*$ is the adjoint of $A$. However, if $W(Au)$ is non-convex,
to find the Legendre dual function is very difficult, or even impossible. For example, if we let $p = Au = u_{xx}$, the stored energy in problem (1), namely

$$W(p - \mu) = \int_0^1 \frac{1}{2} E(\frac{1}{2} p^2 + cp - \mu)^2 dx,$$

is a double-well function for $\mu > 0$. Since the dual variable

$$p^* = DW(p) = (p + c)(\frac{1}{2} p^2 + cp - \mu)$$

is nonlinearly dependent on $p$, the Legendre dual function $W^c(p^*)$ does not have a simple algebraic expression (Sewell 1987). The Fenchel-conjugate function $W^*(p^*)$ is always convex and lower semicontinuous, but if $W(p)$ is non-convex, there exists a dual gap between the primal and dual problems, that is, $\inf J(u, Au) \geq \sup J(A^* p^*, p^*)$. In this case, the dual solutions may be useless (Rockafellar 1994). Thus, the well-developed classical duality theory can be used only for geometrically linear systems with convex stored energy $W(p)$.

In finite deformation theory, the geometrically linear deformation measure $Au = \nabla u$ is not a strain tensor. Its dual variable $p^*$ can be considered as the first Piola–Kirchhoff stress. For most materials, the stored energy $W$ is usually non-convex in the deformation gradient $\nabla u$; see, for example, (Ogden 1984; Truesdell & Noll 1992). However, $W$ is often convex in finite (that is, geometrically nonlinear) strain measures (Hill, 1978). For example, in the non-convex problem (1), if we choose the finite deformation strain $p = Au = \frac{1}{2} u_{xx}^2 + cu_{xx}$, then $W(p)$ is a quadratic function. Its complementary energy $W^c = W^*$ is also a quadratic function and the constitutive relation $p^* = DW(p)$ is linear. Actually, according to Tonti (1972), in many physical theories we can identify some intermediate variables, such as velocity and momentum in dynamical systems, strain and stress in continuum mechanics, and so on. A characteristic of these intermediate variables is that they appear always in pairs (that is, one-to-one). The duality relations between these paired intermediate variables describe the internal (or constitutive) properties of systems.

The key idea of the nonlinear dual transformation method proposed in this paper is to choose the appropriate operator $A$ to ensure that the stored energy $W(p)$ is either convex or concave. Actually, this method has been used by Gao & Strang (1989a,b) in geometrically nonlinear variational/boundary value problems with convex stored energy $W$. In order to recover the lost symmetry between the primal and dual variational problems, they introduced the so-called complementary gap function, which leads to a nonlinear Lagrange duality theory in fully nonlinear variational problems. They proved that if this gap function is positive on the equilibrium admissible field, the system has a unique dual problem. Applications of this general duality theory have been given in a series of publications on finite elastoplasticity (see Gao 1992, 1994, 1995, 1996a, b,c, 1997, 1998a, b, 1999; Gao & Cheung 1989; Gao & Strang 1989a, b; Gao & Yang 1995). Some open problems in large deformation plastic limit analysis were partially solved (Gao 1994; Gao & Strang 1989b). In the study of geometrically nonlinear variational inequality problems, it was realized that by using this nonlinear Lagrange duality theory, the nonlinear von Karman equations can be transformed into a coupled quadratic variational problem (Yau & Gao 1992). Recently, in the study of post-buckling analysis of large deformation beam theory, it is discovered (see Gao 1996b, 1997) that for convex stored energy, if the gap function is negative, the
generalized complementary energy $L(u, p^*)$ is a so-called super-critical point functional, and a pure complementary energy principle (in terms of the conjugate stress only) was established in the extended beam theory. But in general fully nonlinear systems, how to construct a truly dual variational principle remains unclear. Actually, in finite deformation theory, the question concerning the existence of a pure complementary energy variational principle has been argued for more than forty years.

In the present paper, this open problem is solved completely. We propose a reasonably complete complementary-duality theory in general fully nonlinear parametric variational problems. Detailed mathematical proofs are given for the complementary extremum principles recently proposed by the author (Gao 1997). The original motivation for this research was the desire to complete the work of Gao & Strang on duality theory in geometrically nonlinear systems, and to solve the remaining open problems in finite deformation theory. It is now realized that this research is important for problems in phase transitions, post-bifurcation analysis, nonlinear partial differential equations, variational inequalities, non-convex optimization and mathematical programming, and other areas.

The rest of this paper is divided into five main sections. The next section sets up the notation used in the paper and describes the problems. A general framework in fully nonlinear systems and the complementary gap function are discussed. Section 3 presents a Lagrangian duality theory in fully nonlinear systems. The critical points in fully nonlinear systems are classified. Section 4 describes a method of obtaining a truly dual variational principle from a geometrically nonlinear Lagrangian. An interesting triality theory is presented, which contains a minimax dual extremum principle and a pair of minimum and maximum dual extremum principles. Applications in finite deformation theory are illustrated in Section 5 and a pure complementary energy variational principle is constructed. It shows that the dual Euler–Lagrange equation is an algebraic tensor equation. A general analytic solution for mixed boundary value problems in finite deformation theory is obtained. Section 6 shows the applications of this general analytic solution in one-dimensional nonconvex variational problems with concave, double and triple-well energies. The relationship between the algebraic curves and nonlinear differential equations is found. A concluding remark is given in the last section.

This paper is dedicated to the memory of my beloved wife, Professor Rosa Huang (1963–1995). The wonderful duality in my life is broken, the triality theory comes. The intrinsic beauty in triality will tell the true story and the price we paid.

2. Primal problem and framework in fully nonlinear systems

Let $\mathcal{U}$, $\mathcal{U}^*$ and $\mathcal{E}$, $\mathcal{E}^*$ be two pairs of real topological vector spaces, in duality with respect to certain bilinear forms $(\cdot, \cdot) : \mathcal{U} \times \mathcal{U}^* \to \mathbb{R}$ and $(\cdot, \cdot) : \mathcal{E} \times \mathcal{E}^* \to \mathbb{R}$, respectively. Let $\Lambda : \mathcal{U} \to \mathcal{E}$ be a continuous, nonlinear operator. Then the so-called geometrical equation can be written as

$$p = \Lambda u.$$  (2)

We assume that $\Lambda$ is Gâteaux (or G-) differentiable. The directional derivative of $p$ at $u$ in the direction $v \in \mathcal{U}$ is defined as

$$\delta p(u; v) := \lim_{\theta \to 0^+} \frac{p(u + \theta v) - p(u)}{\theta} = \Lambda_t(u)v,$$
where $\Lambda_t$ is the Gâteaux derivative of the operator $A$ at $u$. According to Gao & Strang (1989a), we have the decomposition
\[ A = \Lambda_t + \Lambda_n, \]
(3)
where $\Lambda_n$ is a complementary operator of $\Lambda_t$, which plays an important role in geometrically nonlinear variational principles.

Let $W : E \to \mathbb{R} := [-\infty, +\infty]$ be a real-valued functional, finite and Gâteaux differentiable on a convex subset $D \subset E$. The duality relation between the pairing spaces $E$ and $E^*$ can be described by the so-called constitutive equation
\[ p^* = DW(p), \]
(4)
where $DW : D \subset E \to E^*$ stands for the Gâteaux-derivative of $W$. In conservative systems, we can usually choose a suitable operator $\Lambda$ such that the stored energy $W$ is either convex or concave in $p = \Lambda u$. But $W(\Lambda u)$ may be non-convex in $\Lambda$ because of the nonlinearity of $\Lambda$.

Similarly, by introducing a real valued functional $F : U \to \mathbb{R}$, finite and Gâteaux differentiable on a convex set $C \subset U$, the duality relation between $U$ and its dual space $U^*$ can be given by
\[ u^* = DF(u). \]
(5)
If $F$ is a linear functional then, by the Riesz representation theorem, $F(u)$ can be written as $F(u) = (\bar{u}^*, u)$, where the given dual variable $\bar{u}^* \in U^*$ is the so-called source variable in mathematical physics problems.

For a given parameter $\mu \in \mathcal{E}$, the total potential $P : U \to \mathbb{R}$ is defined as
\[ P_\mu(u) = W(\Lambda u - \mu) - F(u). \]
(6)
Then the inf-primal problem under consideration can be stated as follows.

**Problem 1** For a given $\mu \in \mathcal{E}$, find $\bar{u} \in U$ such that
\[ (P_{\text{inf}}) : \quad P_\mu(\bar{u}) = \inf \ P_\mu(u) \ \forall u \in U. \]
(7)
Note that $P : U \to \mathbb{R}$ is finite, G-differentiable at $u$ if and only if
\[ u \in U_0 := \{ u \in U \mid u \in C, \ \Lambda u \in D \}. \]
(8)
Therefore, we call that $u \in U_0$ the implicit constraint of $(P_{\text{inf}})$, and $U_0$ is the so-called feasible (or kinetically admissible) space. A point $\bar{u} \in U_0$ is called a critical point of $P_\mu$ if $DP_\mu(\bar{u}) = 0$, which leads to the Euler–Lagrange equation (see Gao & Strang 1989a)†
\[ \Lambda_t^*(\bar{u})D_pW(\Lambda\bar{u} - \mu) - DF(\bar{u}) = 0, \]
(9)
† In the proof given in Gao & Strang (1989a), the extremality condition $DP(\bar{u}) = 0$ was considered to be equivalent to the subdifferential condition $0 \in \partial P(\bar{u})$. This is true only when $P$ is G-differentiable and convex. See the following theorem.
where $D_W$ denotes the G-derivative of $W$ with respect to $p = Au$, and the equilibrium operator $\Lambda^*_t : \mathcal{E}^* \rightarrow \mathcal{U}^*$ is the adjoint of $\Lambda_t$ defined by

$$
(p^*, A_t(\bar{u})u) = (\Lambda^*_t(\bar{u})p^*, u) \quad \forall u \in \mathcal{U}_d.
$$

(10)

All the critical points of $P_\mu$ form a subset of $\mathcal{U}_d$, denoted by $\mathcal{U}_c = \{ \bar{u} \in \mathcal{U}_d \mid DP_\mu(\bar{u}) = 0 \}$. All the solutions of $(P_{\text{int}})$ form a subset of $\mathcal{U}_d$, denoted by $\mathcal{U}_s := \{ \bar{u} \in \mathcal{U}_d \mid P_\mu(\bar{u}) = \inf P_\mu(u) \quad \forall u \in \mathcal{U} \}$. According to nonlinear analysis, if $\mathcal{U}_d$ is a non-empty closed and bounded convex subset of a reflexive Banach space $\mathcal{U}$, then the solution set $\mathcal{U}_s$ of $(P_{\text{int}})$ is a non-empty convex subset of $\mathcal{U}_d$. In non-convex problems, a critical point $\bar{u} \in \mathcal{U}_c$ is not necessarily a solution of $(P_{\text{int}})$. It could be either a saddle point or a local extremum of $P_\mu$. However, if $\mathcal{U}_d$ is an open set, a solution $\bar{u} \in \mathcal{U}_s$ should be a critical point of $P_\mu$.

Equation (9) is also known as the fundamental equation of nonlinear equilibrium problems, which can be split as the following three canonical equations:

1. Geometrical equation: $p = Au$;
2. Duality equations: $p^* = DW(p - \mu)$, $u^* = DF(u)$;
3. Equilibrium equation: $u^* = \Lambda^*_t(u)p^*$.

(11)

In mathematical physics, if we call $\mathcal{U}$ the configuration vector space, its dual space $\mathcal{U}^*$ should be the so-called source vector space, $\mathcal{E}$ and its dual $\mathcal{E}^*$ are called the paired intermediate variable spaces. According to Tonti (1972), in most physical theories, the intermediate variables $p$ and $p^*$ appear always in pairs, that is, by choosing a suitable geometric operator $\Lambda : \mathcal{U} \rightarrow \mathcal{E}$, the constitutive equation $p^* = DW(p)$ is one-to-one. So, the geometrical equation (2) is also called the definition equation (cf. Tonti 1972), which does not involve any physical property of the problem.

**Definition 1** Suppose that for a given problem $(P_{\text{int}})$, the geometric operator $\Lambda : \mathcal{U} \rightarrow \mathcal{E}$ can be chosen in such a way that the constitutive relation $DW : \mathcal{D} \subset \mathcal{E} \rightarrow \mathcal{E}^*$ is one-to-one. Then the problem is called geometrically nonlinear if the operator $\Lambda$ is nonlinear; it is called physically nonlinear if at least one of the duality relations is nonlinear; and it is called fully nonlinear if it is both geometrically and physically nonlinear.

Obviously, the problem is linear if it is both geometrically and physically linear. In this case, the equilibrium operator $\Lambda^*$ is simply the adjoint of $\Lambda$. If, for a given $\bar{u}^* \in \mathcal{U}^*$, $F(u) = \langle u, \bar{u}^* \rangle$ is linear and $W(p) = \frac{1}{2}Cp, p$ is quadratic, where $C : \mathcal{E} \rightarrow \mathcal{E}^*$ is a linear operator, then the fundamental equation can be written as

$$
\Lambda^*C Au = \bar{u}^*.
$$

(12)

If $C$ is symmetric, then the operator $K = \Lambda^*CA : \mathcal{U} \rightarrow \mathcal{U}^*$ is self-adjoint $K = K^*$. Further, $K$ is an elliptic operator if $C$ is strictly positive definite; $K$ is hyperbolic if $C$ is indefinite. In the textbook by Strang (1986), this nice symmetrical fundamental equation can be seen in both continuous theories and discrete systems. However, the symmetry is broken in geometrically nonlinear systems. The framework for a fully nonlinear system is shown in Fig. 1.

Using the operator decomposition (3), the relation between the pairing spaces $\mathcal{U}$, $\mathcal{U}^*$ and $\mathcal{E}$, $\mathcal{E}^*$ should be

$$
\langle p^*, p \rangle = (p^*, A_t u) + (p^*, A_n u) = (u^*, u) - G(u, p^*),
$$

(13)
where
\[ G(u, p^*) := (p^*, -\Lambda_n(u)u) \] (14)

A function \( F : \mathcal{U} \to \mathbb{R} \) is lower semicontinuous (l.s.c.) if
\[ \lim_{u_n \to u} \inf F(u_n) \geq F(u) \quad \forall u \in \mathcal{U}. \] (15)
However, \( F : \mathcal{U} \to \mathbb{R} := (-\infty, +\infty) \) is upper semicontinuous (u.s.c.) if \(-F\) is l.s.c. We make the following assumptions for some of our results.

(A1) \begin{align*}
F : \mathcal{U} \to \mathbb{R} & \text{ is concave, u.s.c. and G-differentiable on } \mathcal{C} \subset \mathcal{U}, \\
W : \mathcal{E} \to \mathbb{R} & \text{ is convex, l.s.c. and G-differentiable on } \mathcal{D} \subset \mathcal{E}.
\end{align*}

(A2) \( F : \mathcal{C} \subset \mathcal{U} \to \mathbb{R} \) is linear and \( \Lambda : \mathcal{U} \to \mathcal{E} \) is a quadratic operator.

The following theorem generalizes the result proposed by Gao & Strang (1989a).

THEOREM 1 (Minimum potential energy theorem) Let us assume, in addition to (A1), that \( \Lambda : \mathcal{U} \to \mathcal{E} \) is quadratic, \( \bar{u} \in \mathcal{U}_c \) is a critical point of \( P_\mu \) and that \( \bar{p}^* = DW(\Lambda \bar{u} - \mu) \). If \( G(\bar{u}, \bar{p}^*(\bar{u})) \geq 0 \), then \( \bar{u} \) is a global minimizer of \( P_\mu \). If \( G(\bar{u}, \bar{p}^*(\bar{u})) < 0 \), then \( \bar{u} \in \mathcal{U}_c \) could be either a local minimizer or a local maximizer of \( P_\mu \). If the gap function is positive on \( \mathcal{U}_c \), that is, \( G(\bar{u}, \bar{p}^*(\bar{u})) \geq 0 \quad \forall \bar{u} \in \mathcal{U}_c \), then the solution set \( \mathcal{U}_c \) of \( (P_{inf}) \) is a closed convex subset of \( \mathcal{U}_0 \) and all critical points \( \bar{u} \in \mathcal{U}_c \) are global minimizers of \( P_\mu \). The set \( \mathcal{U}_c \) contains a unique critical point of \( P_\mu \) if the gap function is strictly positive on \( \mathcal{U}_c \).
Proof. Since $W$ and $F$ are G-differentiable on $D$ and $C$, respectively, by the convexity of $W$ and concavity of $F$, we have

\begin{align}
W(Au) - W(Au) & \geq \langle DW(Au), Au - Au \rangle \quad \forall Au \in D, \\
F(u) - F(u) & \leq \langle DF(u), u - u \rangle \quad \forall u \in C.
\end{align}

(16) (17)

Since $A$ is a quadratic operator, for any given $u = \bar{u} + \delta u$, we have (see Gao & Strang, 1989a)

$$A(u)u = A(\bar{u})\bar{u} + A_1(\bar{u})\delta u - A_\mu(\delta u)\delta u.$$  

(18)

If $\bar{u} \in U_\alpha$ is a critical point of $P$, we have

$$P_\mu(u) - P_\mu(\bar{u}) \geq \langle DW_\mu(A\bar{u}), A_1(\bar{u})\delta u \rangle - \langle DW_\mu(A\bar{u}), A_\mu(\delta u)\delta u \rangle - \langle DF(\bar{u}), \delta u \rangle$$

$$= (A^*_1(\bar{u})DW(A\bar{u}) - DF(\bar{u}), \delta u) + G(\delta u, \bar{p}^*(\bar{u}))$$

$$= G(u - \bar{u}, \bar{p}^*) \quad \forall u \in U_\alpha.$$

Since the operator $A : U \to \mathcal{E}$ is quadratic, the gap function $G(u, P^*)$ should be a pure quadratic functional of $u \in U_\alpha$. If $G(\bar{u}, \bar{p}^*) > 0$, then $G(u, \bar{p}^*(\bar{u})) > 0 \quad \forall u \in U_\alpha$. So, $P_\mu(u) - P_\mu(\bar{u}) > 0 \quad U_\alpha \subset U_\alpha$, and $\bar{u}$ is a global minimizer of $P_\mu$. If $G(\bar{u}, \bar{p}^*) < 0$, $P_\mu$ is then non-convex. The proof is given by the Triality theorem I in Section 4. If the gap function is positive on $U_\alpha$, then $P_\mu : U_\alpha \to \mathcal{R}$ is convex. Apart from the trivial cases where the infimum in $(P_{\text{int}})$ (denoted by $\alpha$) is equal to $\pm \infty$, the set $U_\alpha$ of solutions of $(P_{\text{int}})$ is

$$U_\alpha = \{ \bar{u} \in U_\alpha \mid P_\mu(\bar{u}) \leq \alpha \}.$$

Since $P_\mu : U_\alpha \to \mathcal{R}$ is convex, for any given $\theta \in [0, 1],$

$$P_\mu(\theta\bar{u}_1 + (1 - \theta)\bar{u}_2) \leq \theta P_\mu(\bar{u}_1) + (1 - \theta)P_\mu(\bar{u}_2) \leq \alpha \quad \forall \bar{u}_1, \bar{u}_2 \in U_\alpha,$$

that is, $\bar{u}_1 + (1 - \theta)\bar{u}_2 \in U_\alpha$, $\forall \bar{u}_1, \bar{u}_2 \in U_\alpha$. So $U_\alpha$ is convex. It is closed since $P_\mu$ is l.s.c. (cf. Ekeland & Temam 1976). If the gap function is strictly positive on $U_\alpha$, $P_\mu$ is strictly convex and, hence, $P_\mu$ has at most one critical point. \hfill \square

**Corollary 1** In addition to (A2), we assume that $W : \mathcal{E} \to \mathcal{R}$ is concave. Then, for a given $\bar{u} \in U_\alpha$, and associated $\bar{p}^*(\bar{u})$, if $G(u, \bar{p}^*(\bar{u})) \leq 0 \quad \forall u \in U_\alpha$, then $\bar{u}$ is a global maximizer of $P_\mu$.

The proof of this corollary is similar to the proof of Theorem 1.

### 3. Geometrically nonlinear Lagrangian and critical points

To establish the dual problem, we need to find the complementary energy of the system. For a given functional $W : \mathcal{E} \to \mathcal{R}$, its conjugate functional is defined by the Fenchel transformation

$$W^*(p^*) := \sup_{p \in \mathcal{E}} \{ (p^*, p) - W(p) \}.$$  

(19)
which is always convex and l.s.c. If \( W \) is strictly convex, \( G \)-differentiable on \( D \), then (19) is the classical Legendre transformation and the following relations are equivalent to each other:

\[
p^* = DW(p) \iff p = DW^*(p^*) \iff W(p) + W^*(p^*) = \langle p, p^* \rangle.
\]

(20)

By letting \( q = p - \mu \), the conjugate functional of \( W_\mu(p) = W(p - \mu) \) is

\[
W^*_\mu(p^*) = \sup_q \{ \langle p^*, q + \mu \rangle - W(q) \}
= W^*(p^*) + \langle p^*, \mu \rangle.
\]

(21)

For a concave functional \( F : \mathcal{U} \to \mathbb{R} \), its conjugate function should be

\[
F^*(u^*) = \inf_{u \in \mathcal{U}} \{ \langle u^*, u \rangle - F(u) \},
\]

(22)

which is always concave and u.s.c.

Let \( \mathcal{C}^* \subset \mathcal{U}^* \) and \( \mathcal{D}^* \subset \mathcal{E}^* \) be non-empty, convex subsets on which \( F^* \) and \( W^*_\mu \) are finite and \( G \)-differentiable, respectively. The total complementary energy for geometrically nonlinear problems is (see Gao & Strang 1989a)

\[
P^*_\mu(p^*, u) := F^*(A_t^*(u)p^*) - W^*_\mu(p^*) - G(u, p^*).
\]

(23)

It may be that \( P^*_\mu : \mathcal{E}^* \times \mathcal{U} \to \mathbb{R} \) is neither concave nor convex. It is finite and \( G \)-differentiable on the so-called equilibrium admissible space defined as

\[
\mathcal{E}^*_u := \{(u, p^*) \in \mathcal{C} \times \mathcal{D}^* | A_t^*(u)p^* = DF(u)\}.
\]

(24)

The following complementary energy theorem was proved in Gao & Strang, (1989a).

THEOREM 2 (Maximum Complementary Energy Theorem) Under assumption (A2), if \( W : \mathcal{E} \to \mathbb{R} \) is convex and

\[
G(u, p^*) \geq 0 \quad \forall (u, p^*) \in \mathcal{E}^*_u,
\]

(25)

then the complementary energy variational problem

\[
(P^*_\text{sup}) : \sup_{(u, p^*) \in \mathcal{E}^*_u} P^*_\mu(p^*, u)
\]

(26)

is equivalent to the primal problem \((P_{\text{inf}})\) in the sense that they have same solution set and

\[
\inf_{u \in \mathcal{U}} P^*_\mu(u) = \sup_{(u, p^*) \in \mathcal{E}^*_u} P^*_\mu(p^*, u).
\]

(27)

In the problem \((P^*_\text{sup})\), the variable \( u \) depends on \( p^* \) through the equilibrium equation \( A_t^*(u)p^* = DF(u) \). If we can write \( u \) in terms of \( p^* \), then the pure complementary energy should be

\[
P^*_u(p^*) = P^*_\mu(p^*, u(p^*)).
\]

(28)

We will see that this pure complementary energy is actually a truly dual functional of \( P_\mu \) and can be obtained through the following geometrically nonlinear Lagrangian.
The Lagrangian form \( L : U \times E^* \to \mathbb{R} \) in geometrically nonlinear systems was proposed in Gao & Strang (1989a) as

\[
L_\mu(u, p^*) := \langle Au, p^* \rangle - W_\mu^*(p^*) - F(u),
\]

which is finite and G-differentiable on \( C \times D^* \). A point \((\bar{u}, \bar{p}^*) \in C \times D^*\) is said to be a critical point of \( L_\mu \) if

\[
D_u L_\mu(\bar{u}, \bar{p}^*) = 0, \quad D_p L_\mu(\bar{u}, \bar{p}^*) = 0.
\]

Here \( D_u, D_p \) denote the partial Gateaux derivatives on \( U \) and \( E^* \), respectively. It is easy to show that

\[
D_u L_\mu(\bar{u}, \bar{p}^*) = 0 \Rightarrow \Lambda^*_\mu(\bar{u}) \bar{p}^* - DF(\bar{u}) = 0,
\]

\[
D_p L_\mu(\bar{u}, \bar{p}^*) = 0 \Rightarrow \Lambda \bar{u} - DW_\mu^*(\bar{p}^*) = 0.
\]

In fully nonlinear systems, we need the following definition.

**DEFINITION 2** A point \((\bar{u}, \bar{p}^*)\) is said to be a **right saddle point** of \( L_\mu \) if

\[
L_\mu(u, \bar{p}^*) \geq L_\mu(\bar{u}, \bar{p}^*) \geq L_\mu(u, p^*) \quad \forall (u, p^*) \in U \times E^*.
\]

A point \((\bar{u}, \bar{p}^*)\) is said to be a **left saddle point** of \( L_\mu \) if

\[
L_\mu(u, \bar{p}^*) \leq L_\mu(\bar{u}, \bar{p}^*) \leq L_\mu(u, p^*) \quad \forall (u, p^*) \in U \times E^*.
\]

A point \((\bar{u}, \bar{p}^*)\) is said to be a **sub-critical** (or \( \partial^- \)-critical) point of \( L_\mu \) if

\[
L_\mu(\bar{u}, p^*) \geq L_\mu(\bar{u}, \bar{p}^*) \leq L_\mu(u, p^*) \quad \forall (u, p^*) \in U \times E^*.
\]

A point \((\bar{u}, \bar{p}^*)\) is said to be a **super-critical** (or \( \partial^+ \)-critical) point of \( L_\mu \) if

\[
L_\mu(\bar{u}, p^*) \leq L_\mu(\bar{u}, \bar{p}^*) \geq L_\mu(u, \bar{p}^*) \quad \forall (u, \bar{p}^*) \in U \times E^*.
\]

Obviously, the left saddle point (respectively sub-critical point) of \( L_\mu \) should be a right saddle point (respectively super-critical point) of \(-L_\mu\). A right saddle point is also simply called the **saddle point**. According to non-smooth analysis, for any given real valued function \( F : U \to \mathbb{R} \), the sub-differential and super-differential of \( F \) at \( u \in U \) are defined by

\[
\partial^- F(u) = \{ u^* \in U^* \mid F(v) - F(u) \geq \langle u^*, v - u \rangle \ \forall v \in U \},
\]

\[
\partial^+ F(u) = \{ u^* \in U^* \mid F(v) - F(u) \leq \langle u^*, v - u \rangle \ \forall v \in U \},
\]

respectively. In convex analysis, \( \partial^- \) is simply written as \( \partial \). The super-differential \( \partial^+ \) is also called the **over-differential**, written as \( \partial^- \) (see Aubin & Ekeland, 1984). It is easy to check that \( \partial^+ F = -\partial^- (-F) \). Moreover, if \( F \) is G-differentiable on \( C \), then

\[
\partial^+ F(u) = -\partial^- (-F(u)) = \{ DF(u) \} \ \forall u \in C.
\]

So, the inequalities in (35) are equivalent to the partial sub-differentials of \( L_\mu \), that is

\[
0 \in \partial^- u L_\mu(\bar{u}, \bar{p}^*), \quad 0 \in \partial^+ p L_\mu(\bar{u}, \bar{p}^*).
\]
In geometrically linear systems, this is Auchmuty’s definition of $\partial$-critical point (Auchmuty, 1983). The inequalities in (36) are equivalent to the partial super-differentials of $L_\mu$, so that

$$0 \in \partial_{\tilde{\mu}}^+ L_\mu(\tilde{u}, \tilde{p}^\ast), \quad 0 \in \partial_{\tilde{\mu}}^+ L_\mu(\tilde{u}, \tilde{p}^\ast).$$

By the Legendre transformation, the Hamiltonian in fully nonlinear systems can be obtained from the Lagrangian (see Gao & Strang 1989a) as

$$H_\mu(u, p^\ast) = \langle Au, p^\ast \rangle - L_\mu(u, p^\ast). \quad (37)$$

If $W$ and $F$ are convex, then the three canonical equations (11) can be written (Gao & Strang 1989a) in the forms

$$\Lambda \tilde{u} \in \partial_{\tilde{p}^\ast} H_\mu(\tilde{u}, \tilde{p}^\ast), \quad \Lambda^* \tilde{p}^\ast \in \partial_{\tilde{u}} H_\mu(\tilde{u}, \tilde{p}^\ast). \quad (38)$$

If $\Lambda$ is a linear operator, then the inequalities (36) are equivalent to symmetric canonical forms in geometrically linear systems, namely

$$\Lambda \tilde{u} \in \partial_{\tilde{p}^\ast} H_\mu(\tilde{u}, \tilde{p}^\ast), \quad \Lambda^* \tilde{p}^\ast \in \partial_{\tilde{u}} H_\mu(\tilde{u}, \tilde{p}^\ast),$$

which corresponding to Auchmuty’s definition of anomalous critical point for convex $W$ and $F$. In dynamical systems $\Lambda = d/dt$ is a linear operator. Its adjoint is $\Lambda^* = -d/dt$, and $G(u, p^\ast) = 0$. If $W(Au) = \frac{1}{2} \langle Au, CAu \rangle$ is quadratic and the operator $K = \Lambda^* C \Lambda = K^*$ is self-adjoint, then the total potential $P_\mu$ is the so-called total action, written as

$$I(u) = \frac{1}{2} \langle u, Ku \rangle - F(u). \quad (39)$$

Replacing $p^\ast$ by $CAu$, the complementary energy $-P^c$ in this case is the well-known Clarke dual action (Clarke 1979, 1985), denoted by

$$I^c(u) = \frac{1}{2} \langle u, Ku \rangle - F^c(Ku). \quad (40)$$

In fully nonlinear systems, the symmetry in canonical forms is lost. The properties of critical points of $L_\mu$ depend on the gap function; see (Gao 1997).

**THEOREM 3** (Critical point theorem) Suppose that the assumption (A2) holds and $(\tilde{u}, \tilde{p}^\ast)$ is a critical point of $L_\mu$. If $W : \mathcal{D} \subset \mathcal{E} \rightarrow \mathbb{R}$ is convex and G-differentiable, then

$$(\tilde{u}, \tilde{p}^\ast)$$

is a right saddle point of $L_\mu$ if and only if $G(u, \tilde{p}^\ast) \geq 0 \ \forall u \in \mathcal{C}$,

$$(\tilde{u}, \tilde{p}^\ast)$$

is a super-critical point of $L_\mu$ if and only if $G(u, \tilde{p}^\ast) < 0 \ \forall u \in \mathcal{C}. \quad (41)$$

If $W : \mathcal{D} \subset \mathcal{E} \rightarrow \mathbb{R}$ is concave and G-differentiable, then

$$(\tilde{u}, \tilde{p}^\ast)$$

is a left saddle point of $L_\mu$ if and only if $G(u, \tilde{p}^\ast) \leq 0 \ \forall u \in \mathcal{C}$,

$$(\tilde{u}, \tilde{p}^\ast)$$

is a sub-critical point of $L_\mu$ if and only if $G(u, \tilde{p}^\ast) > 0 \ \forall u \in \mathcal{C}. \quad (42)$$

**Proof.** Since $F$ is linear, $F(\tilde{u} + \delta u) = F(\tilde{u}) + \langle DF(\tilde{u}), \delta u \rangle$. Considering (18), we have

$$L_\mu(u, \tilde{p}^\ast) - L_\mu(\tilde{u}, \tilde{p}^\ast) = \langle \Lambda^*_{\tilde{u}}(\tilde{u}) \tilde{p}^\ast - DF(\tilde{u}), \delta u \rangle + G(\delta u, \tilde{p}^\ast)$$

$$= \{D_u L_\mu(\tilde{u}, \tilde{p}^\ast), \delta u \} + G(\delta u, \tilde{p}^\ast).$$
Noting that \((\bar{u}, \bar{p}^*)\) is a critical point of \(L_\mu\), so that \(D_u L_\mu(\bar{u}, \bar{p}^*) = 0\), we then have
\[
L_\mu(u, \bar{p}^*) - L_\mu(\bar{u}, \bar{p}^*) = G(u - \bar{u}, \bar{p}^*),
\]
and hence
\[
L_\mu(u, \bar{p}^*) \geq L_\mu(\bar{u}, \bar{p}^*) \quad \text{if and only if} \quad G(u, \bar{p}^*) \geq 0 \quad \forall u \in \mathcal{U},
\]
\[
L_\mu(u, \bar{p}^*) \leq L_\mu(\bar{u}, \bar{p}^*) \quad \text{if and only if} \quad G(u, \bar{p}^*) \leq 0 \quad \forall u \in \mathcal{U}.
\]
(43)

If \(W : \mathcal{D} \subset \mathcal{E} \to \mathbb{R}\) is convex and G-differentiable, then \(L_\mu : \mathcal{E}^* \to \mathbb{R}\) is concave for any given \(u \in \mathcal{U}\). If \((\bar{u}, \bar{p}^*)\) is a critical point of \(L_\mu\), we have
\[
L_\mu(u, p^*) \leq L_\mu(\bar{u}, \bar{p}^*) \quad \forall p^* \in \mathcal{E}^*.
\]
The statement (41) is proved by combining this with (43).

However, if \(W : \mathcal{D} \subset \mathcal{E} \to \mathbb{R}\) is concave and G-differentiable, then \(L_\mu : \mathcal{E}^* \to \mathbb{R}\) is convex for any given \(u \in \mathcal{U}\). If \((\bar{u}, \bar{p}^*)\) is a critical point of \(L_\mu\), we have
\[
L_\mu(u, p^*) \geq L_\mu(\bar{u}, \bar{p}^*) \quad \forall p^* \in \mathcal{E}^*.
\]
The statement (42) is then proved by combining this with (43). 

From Theorem 3, the following corollary can easily be obtained for convex \(W\).

**Corollary 2** Suppose that (A2) holds and \(W : \mathcal{E} \to \mathbb{R}\) is convex. The gap function \(G(u, p^*)\) is positive on \(\mathcal{E}^*_u\) if and only if all critical points of \(L_\mu\) are saddle points. Moreover, if the gap function is strictly positive on \(\mathcal{E}^*_u\), and \(W^* : \mathcal{E}^* \to \mathbb{R}\) is strictly convex, then \(L_\mu\) has at most one saddle point.

In fully nonlinear problems, the total potential energy \(P_\mu\) is usually non-convex. Then \(P_\mu\) may have a local maximizer on a subset of \(\mathcal{U}\), which could be a critical point in phase transitions. So on \(\mathcal{U}_a \subset \mathcal{U}\), we can propose the following sup-primal problem.

**Problem 2** For a given \(\mu \in \mathcal{E}\), find \(\bar{u} \in \mathcal{U}_a\) such that
\[
(P_{\text{sup}}) : \quad P_\mu(\bar{u}) = \sup_{u \in \mathcal{U}_a} P_\mu(u) \quad \forall u \in \mathcal{U}_a.
\]
(46)

The relationship between the potential energy and the Lagrangian depends on the convexity of \(W\).

**Proposition 1** For any given function \(F : \mathcal{U} \to \mathbb{R}\),
\[
P_\mu(u) = \begin{cases} 
\sup_{p^* \in \mathcal{E}^*} L_\mu(u, p^*) & \text{if } W : \mathcal{E} \to \mathbb{R} \text{ is convex,} \\
\inf_{p^* \in \mathcal{E}^*} L_\mu(u, p^*) & \text{if } W : \mathcal{E} \to \mathbb{R} \text{ is concave.}
\end{cases}
\]
(47)

This proposition can easily be proved by using the Fenchel transformation. The relationship between the complementary energy \(P_\mu^*\) and \(L_\mu\) depends on the gap function. Generally speaking, for any given \(p^* \in \mathcal{E}^*\) such that \(\mathcal{E}^*_u\) is not empty, the pure complementary energy functional can be defined by
\[
P_\mu^*(p^*) = \sup_{u \in \bar{U}} L_\mu(u, p^*).
\]
4. Complementary extremum principles

In this section, we assume that conditions in (A2) hold. The following theorem shows how to find the pure complementary energy $P^*_\mu$.

**THEOREM 4** (Pure complementary energy theorem) For a given $p^* \in \mathcal{E}^*$ such that $\mathcal{E}_u^*$ is not empty, we have

$$P^*_\mu(p^*) = \begin{cases} \inf_{u \in \mathcal{U}} L(u, p^*) & \text{if } G(u, p^*) \geq 0 \ \forall u \in \mathcal{U}, \\ \sup_{u \in \mathcal{U}} L(u, p^*) & \text{if } G(u, p^*) < 0 \ \forall u \in \mathcal{U}. \end{cases} \quad (48)$$

**Proof.** Since, for a given $p^* \in \mathcal{E}^*$, the equilibrium admissible space $\mathcal{E}_u^*$ is not empty, there exists at least one $\bar{u} \in \mathcal{U}$ such that $D_u L_\mu(\bar{u}, p^*) = 0$. If $G(u, p^*) \geq 0 \ \forall u \in \mathcal{U}$, then $L_\mu : \mathcal{U} \to \mathbb{R}$ is convex. Thus $\bar{u}$ is a solution of $\inf_u L_\mu(u, p^*)$ and $P^*_\mu(p^*) = L_\mu(\bar{u}, p^*) = \inf_u L_\mu(u, p^*)$. If $G(u, p^*) \leq 0 \ \forall u \in \mathcal{U}$, then $L_\mu : \mathcal{U} \to \mathbb{R}$ is concave. In this case, $\bar{u}$ is a solution of $\sup_u L_\mu(u, p^*)$ and $P^*_\mu(p^*) = \sup_u L_\mu(u, p^*)$. \hfill \Box

The pure complementary energy $P^*_\mu : \mathcal{E}^* \to \mathbb{R}$ is also non-convex, so the following sup- and inf-dual problems can be proposed.

**PROBLEM 3** For a given $\mu \in \mathcal{E}$, find $\bar{p}^*$ such that either

$$(P_{\text{sup}}^*): \quad P^*_\mu(\bar{p}^*) = \sup P^*_\mu(p^*) \ \forall p^* \in \mathcal{E}^*. \quad (49)$$

or

$$(P_{\text{inf}}^*): \quad P^*_\mu(\bar{p}^*) = \inf P^*_\mu(p^*) \ \forall p^* \in \mathcal{E}^*. \quad (50)$$

Let $\mathcal{E}_u^* \subset \mathcal{D}^*$ be a convex subset such that for any given $p^* \in \mathcal{E}_u^*$, the equilibrium admissible space $\mathcal{E}_u^*$ is not empty. Then, on $\mathcal{E}_u^*$, the complementary energy $P^*_\mu$ is finite and can be written as

$$P^*_\mu(p^*) = F^*(\Lambda^*_u p^*) - W^*_\mu(p^*) - G^*(p^*), \quad (51)$$

where

$$G^*(p^*) := G(\bar{u}, p^*) = \langle -\Lambda_u(\bar{u})\bar{u}, p^* \rangle, \quad (52)$$

and $\bar{u}$ is a solution of $D_u L_\mu(\bar{u}, p^*) = 0$ for a given $p^* \in \mathcal{E}_u^*$. The condition $p^* \in \mathcal{E}_u^*$ is called the dual implicit constraint of the sup- and inf-dual problems.

The following lemma is fundamental for our complementary extremum principles.

**LEMMA 1** If $(\bar{u}, \bar{p}^*)$ is any one of the right or left saddle points, the super- or sub-critical points of $L_\mu$, and $L_\mu$ is partially Gâteaux differentiable at $(\bar{u}, \bar{p}^*)$, then $(\bar{u}, \bar{p}^*)$ must be a critical point of $L_\mu$. If $P_\mu$ and $P^*_\mu$ are Gâteaux differentiable at $\bar{u}$ and $\bar{p}^*$, respectively, then $DP_\mu(\bar{u}) = 0$, $DP^*_\mu(\bar{p}^*) = 0$, and

$$P_\mu(\bar{u}) = L_\mu(\bar{u}, \bar{p}^*) = P^*_\mu(\bar{p}^*). \quad (53)$$
Proof. From the theory of convex analysis (see Ekeland & Temam 1976), we know that for any given Gâteaux differentiable functional $F : C \subset \mathcal{U} \to \mathbb{R}$, if either $\bar{u}^* \in \partial^- F(\bar{u})$ or $\bar{u}^* \in \partial^+ F(\bar{u})$ then $\bar{u}^* = DF(\bar{u})$. By Theorem 3, for convex $W$, $(\bar{u}, \bar{p}^*)$ is a saddle-point of $L_\mu$ if and only if $0 \in \partial^- L_\mu(\bar{u}, \bar{p}^*)$ and $0 \in \partial^+ L_\mu(\bar{u}, \bar{p}^*)$. Since $L_\mu$ is partially Gâteaux differentiable at $(\bar{u}, \bar{p}^*)$, we have

$$\partial^- L_\mu(\bar{u}, \bar{p}^*) = \{D \mu L_\mu(\bar{u}, \bar{p}^*)\} = 0, \quad \partial^+ L_\mu(\bar{u}, \bar{p}^*) = \{D \mu^* L_\mu(\bar{u}, \bar{p}^*)\} = 0.$$  

Thus, $(\bar{u}, \bar{p}^*)$ is a critical point of $L_\mu$. Similar results hold for super- and sub-critical points.

Suppose that $(\bar{u}, \bar{p}^*)$ is a saddle point of $L_\mu$. By Theorem 3, $G(\bar{u}, \bar{p}^*)$ is positive for convex $W$. From the definition of the saddle point and Theorem 4, on $\mathcal{U} \times \mathcal{E}^*$ we have

$$P_\mu(\bar{u}) = \sup_{p^*} L_\mu(\bar{u}, p^*) \leq L_\mu(\bar{u}, \bar{p}^*) \leq \inf_u L_\mu(u, \bar{p}^*) = P_\mu^*(\bar{p}^*). \quad (54)$$

On the other hand, for any $u \in \mathcal{U}$ and $p^* \in \mathcal{E}^*$,

$$P_\mu(\bar{u}) - P_\mu^*(\bar{p}^*) = \sup_{p^*} L_\mu(\bar{u}, p^*) - \inf_u L_\mu(u, \bar{p}^*)$$

$$= \sup_{p^*} \sup_u [L_\mu(\bar{u}, p^*) - L_\mu(u, \bar{p}^*)] \geq 0.$$

Hence

$$P_\mu(\bar{u}) = L_\mu(\bar{u}, \bar{p}^*) = P_\mu^*(\bar{p}^*).$$

Also, from

$$P_\mu_u = \sup_{p^*} L_u(\bar{u}, p^*) = L_u(\bar{u}, \bar{p}^*) \geq L_u(\bar{u}, \bar{p}^*) = P_\mu(\bar{u}) \quad \forall u \in \mathcal{U}$$

we have $0 \in \partial^- P_\mu(\bar{u})$. Since $P_\mu$ is Gâteaux differentiable at $\bar{u}$, we have $DP_\mu(\bar{u}) = 0$. Similar results hold for $DP_\mu^*(\bar{p}^*) = 0$ and for the left saddle point.

If $(\bar{u}, \bar{p}^*)$ is a super-critical point of $L_\mu$, then by the definition

$$L_\mu_u(\bar{u}, \bar{p}^*) \leq L_\mu_u(\bar{u}, \bar{p}^*) \geq L_\mu_u(\bar{u}, \bar{p}^*) \quad \forall u \in \mathcal{U}, \quad p^* \in \mathcal{E}^*,$$

and by Theorem 3, we know that $W$ is convex and $G$ is negative. Thus,

$$P_\mu(\bar{u}) = \sup_{p^*} L_\mu(\bar{u}, p^*) = L_\mu(\bar{u}, \bar{p}^*) = \sup_u L_\mu(u, \bar{p}^*) = P_\mu^*(\bar{p}^*),$$

since $(\bar{u}, \bar{p}^*) \in \mathcal{E}^*_u$, and hence $P_\mu^*(\bar{p}^*) = P_\mu(\bar{u})$. Moreover, from

$$P_\mu^*(p^*) = \sup_u L_\mu(u, p^*) = L_\mu(\bar{u}, p^*) \leq L_\mu(\bar{u}, \bar{p}^*) = P_\mu^*(\bar{p}^*),$$

we have $0 \in \partial^+ P_\mu^*(\bar{p}^*)$. The Gâteaux differentiability of $P_\mu^*$ implies that $\bar{p}^*$ is a critical point. Similar results hold for $DP_\mu(\bar{u}) = 0$ and for the sub-critical point. \hfill \square

Based on this lemma, the following complementary extremum principles can be proposed.
THEOREM 5 (Mini-max duality theorem) Suppose that $\mathcal{E}_a^*$ is not empty. If $W : \mathcal{E} \rightarrow \mathbb{R}$ is convex and the gap function $G(u, p^*) = 0 \forall (u, p^*) \in \mathcal{E}_a^*$, then $\bar{u}, \bar{p}^* \in \mathcal{E}_a^*$ is a critical point of $L_\mu$ if and only if

$$P_\mu(\bar{u}) = \inf_{u \in \mathcal{U}} P_\mu(u) = \sup_{p^* \in \mathcal{E}_a^*} P_\mu^*(p^*) = P_\mu^*(\bar{p}^*). \quad (55)$$

If $W : \mathcal{E} \rightarrow \mathbb{R}$ is concave and the gap function $G(u, p^*) = 0 \forall (u, p^*) \in \mathcal{E}_a^*$, then $\bar{u}, \bar{p}^* \in \mathcal{U} \times \mathcal{E}_a^*$ is a critical point of $L_\mu$ if and only if

$$P_\mu(\bar{u}) = \sup_{u \in \mathcal{U}} P_\mu(u) = \inf_{p^* \in \mathcal{E}_a^*} P_\mu^*(p^*) = P_\mu^*(\bar{p}^*). \quad (56)$$

**Proof.** By Corollary 2, for convex $W$ the critical point $(\bar{u}, \bar{p}^*)$ should be a saddle point of $L_\mu$. For any given $u \in \mathcal{U}$, $L_\mu : \mathcal{E}_a^* \rightarrow \mathbb{R}$ is concave, and

$$P_\mu(u) = \sup_{p^* \in \mathcal{E}_a^*} L_\mu(u, p^*) = L_\mu(\bar{u}, \bar{p}^*) \forall u \in \mathcal{U}. \quad \text{By Lemma 1, } L_\mu(\bar{u}, \bar{p}^*) \leq P_\mu(u) \forall u \in \mathcal{U}. \quad \text{Thus } \bar{u} \text{ is a solution of (Pinf). Since the gap function is positive on } \mathcal{E}_a^*, \text{ } L_\mu : \mathcal{U} \rightarrow \mathbb{R} \text{ is convex, then, by Theorem 4,}$

$$P_\mu^*(p^*) = \inf_u L_\mu(u, p^*) = L_\mu(\bar{u}, \bar{p}^*) \leq \inf_p \in \mathcal{E}_a^*.$$

This shows that $\bar{p}^*$ is a solution of $(P_{sup})$ and it follows from Lemma 1 that $P_\mu(\bar{u}) = P_{sup}(\bar{p}^*)$.

Conversely, since $\mathcal{E}_a^* \neq \emptyset$, if $P_\mu(\bar{u}) = \inf P_\mu(u) = \sup P_\mu^*(p^*) = P_{sup}(\bar{p}^*)$ we have, from (47) and (48),

$$P_\mu(\bar{u}) = \sup L_\mu(u, p^*) \geq L_\mu(\bar{u}, p^*) \forall p^* \in \mathcal{E}^*, \quad P_{sup}(\bar{p}^*) = \inf L_\mu(u, \bar{p}^*) \leq L_\mu(u, \bar{p}^*) \forall u \in \mathcal{U}.$$

So the equality $P_\mu(\bar{u}) = P_{sup}(\bar{p}^*)$ shows that $(\bar{u}, \bar{p}^*)$ is a saddle point of $L_\mu$. By Lemma 1, it must be a critical point of $L_\mu$. For concave $W$ and negative gap function the proof is similar. \hfill \square

This theorem shows that for a quadratic operator $\Lambda$, if $W : \mathcal{E} \rightarrow \mathbb{R}$ is convex and the gap function is positive on $\mathcal{E}_a^*$, $P_\mu$ is convex and $P_{sup}$ is concave. The system has only one potential extremum principle and only one complementary extremum principle.\(^\dagger\) However, if the gap function is negative on $\mathcal{E}_a^*$, the system may have more than one set of primal-dual problems.

For a given critical point $(\bar{u}, \bar{p}^*)$ of $L_\mu$, we let $\mathcal{U}_b \times \mathcal{E}_a^* \subset \mathcal{U} \times \mathcal{E}^*$ be its neighbourhood such that on $\mathcal{U}_b \times \mathcal{E}_a^*$, $(\bar{u}, \bar{p}^*)$ is the only critical point of $L_\mu$. Then, we have the following results.

\(^\dagger\) The multi-duality for different $\Lambda$ was discussed in Gao (1992) for finite deformation theory and in Gao & Yang (1995) for minimal surface-type problems.
THEOREM 6 (Maximum duality theorem) If $W : \mathcal{E} \to \mathbb{R}$ is convex and the gap function $G(u, p^*) < 0 \forall (u, p^*) \in \mathcal{E}_u^*$. Then, on $U_0 \times \mathcal{E}_b^*$, sup $P_\mu(u) = \sup P^*_\mu(p^*)$. Moreover, the critical point $(\bar{u}, \bar{p}^*)$ of $L_\mu$ maximizes $L_\mu$ on $U_0 \times \mathcal{E}_b^*$ if and only if

$$P_\mu(\bar{u}) = \sup P_\mu(u) = \sup_{p^* \in \mathcal{E}_b^*} P^*_\mu(p^*) = P^*_\mu(\bar{p}^*).$$  \hspace{1cm} (57)

If $W : \mathcal{E} \to \mathbb{R}$ is concave, $(\bar{u}, \bar{p}^*)$ is a critical point of $L_\mu$, and $G(\bar{u}, \bar{p}^*) \geq 0$, then $(\bar{u}, \bar{p}^*)$ maximizes $L_\mu$ on $U_0 \times \mathcal{E}_b^*$ in either order (that is, either sup$_u$ inf$_p$ $L_\mu$ or sup$_p$ inf$_u$ $L_\mu$) if and only if statement (57) holds.

Proof. By the definition of $L_\mu$ and Lemma 1, for a negative gap function on $\mathcal{E}_u^*$,

$$\sup_{\mu(\Lambda)} P_\mu(u) = \sup_{\mu(\Lambda)} L_\mu(u, p^*) = \sup_{\mu(\Lambda)} P^*_\mu(p^*).$$  \hspace{1cm} (58)

Thus, sup $P = \sup P^*$ since we can take the supremum in either order on $U_0 \times \mathcal{E}_b^*$.

Suppose that $(\bar{u}, \bar{p}^*)$ maximizes $L_\mu$ on $U_0 \times \mathcal{E}_b^*$. Since $U_0 \times \mathcal{E}_b^*$ is a neighbourhood of the critical point $(\bar{u}, \bar{p}^*)$, there exists at least one point $(u, p^*) \in U_0 \times \mathcal{E}_b^*$ such that $W_\mu(\Lambda u) = \sup_{\mu(\Lambda)} (\Lambda u, p^*) - W^*_\mu(p^*)$ is finite. So, on $U_0 \times \mathcal{E}_b^*$, we have

$$L_\mu(\bar{u}, \bar{p}^*) = \sup_{\mu(\Lambda)} L_\mu(u, p^*) = \sup_{\mu(\Lambda)} P_\mu(u).$$  \hspace{1cm} (59)

Then, Lemma 1 gives sup $P_\mu(u) = P_\mu(\bar{u}) = L_\mu(\bar{u}, \bar{p}^*) = P^*_\mu(\bar{p}^*) = \sup P^*_\mu(p^*)$ on $U_0 \times \mathcal{E}_b^*$.

Conversely, if $P_\mu(\bar{u}) = \sup P_\mu(u) = P^*_\mu(p^*) = \sup P^*_\mu(p^*)$ then, for any given $(u, p^*) \in U_0 \times \mathcal{E}_b^*$, such that the gap function is negative, we have

$$P_\mu(\bar{u}) \geq P_\mu(u) = \sup_{p^* \in \mathcal{E}_b^*} L_\mu(u, p^*) \geq L_\mu(\bar{u}, \bar{p}^*),$$

$$P^*_\mu(\bar{p}^*) \geq P^*_\mu(p^*) = \sup_{u \in U_0} L_\mu(u, p^*) \geq L_\mu(\bar{u}, p^*).$$

Thus, by Lemma 1, on $U_0 \times \mathcal{E}_b^*$,

$$L_\mu(\bar{u}, p^*) \leq P^*_\mu(\bar{p}^*) = L_\mu(\bar{u}, \bar{p}^*) = P^*_\mu(\bar{p}^*) \geq L_\mu(\bar{u}, \bar{p}^*).$$

So $(\bar{u}, \bar{p}^*)$ maximizes $L_\mu$ on $U_0 \times \mathcal{E}_b^*$ as required.

For concave $W$, the statement can be proved by a similar method to that given in the proof of the next theorem. \hfill \square

THEOREM 7 (Minimum duality theorem) If $W : \mathcal{E} \to \mathbb{R}$ is convex, $(\bar{u}, \bar{p}^*)$ is a critical point of $L_\mu$, and $G(\bar{u}, \bar{p}^*) < 0$, then $(\bar{u}, \bar{p}^*)$ minimizes $L_\mu$ on $U_0 \times \mathcal{E}_b^*$ in either order (that is, either inf$_u$ sup$_p$ $L_\mu$ or inf$_p$ sup$_u$ $L_\mu$) if and only if

$$P_\mu(\bar{u}) = \inf_{u \in U_0} P_\mu(u) = \inf_{p^* \in \mathcal{E}_b^*} P^*_\mu(p^*) = P^*_\mu(\bar{p}^*).$$  \hspace{1cm} (60)
Proof. From Theorem 3, if $W$ is convex and $G(\bar{u}, \bar{p}^*) < 0$, then $P_\mu(\bar{u}) = \inf P_\mu(a)$. Moreover, the critical point $(\bar{u}, \bar{p}^*)$ of $L_\mu$ minimizes $L_\mu$ on $U_0 \times E^*_b$ if and only if the statement (60) holds.

By Lemma 1, we have $P_\mu(\bar{u}) = L_\mu(\bar{u}, \bar{p}^*) = \inf P_\mu(u)$. Since $\bar{p}^*$ is a critical point of $P_\mu^{*}$ on the open domain $E^*_b$, it should be either a local extremum point or a local saddle point of $P_\mu^{*}$. If $\bar{p}^*$ maximizes $P_\mu^{*}$ on $E^*_b$, then

$$ P_\mu^{*}(\bar{p}^*) = \sup P_\mu^{*}(p^*) = \sup_{u} \inf_{a} L_{\mu}(u, p^*) $$

$$ = \sup_{u} \inf_{a} L_{\mu}(u, p^*) = \sup P_\mu^{*}(u). \quad (61) $$

But $P_\mu^{*}(\bar{p}^*) = P_\mu^{*}(\bar{u})$, and $\bar{u}$ minimizes $P_\mu$ on $U_0$. So this contradiction shows that $\bar{p}^*$ cannot be a local maximizer of $P_\mu^{*}$. If $\bar{p}^*$ is a saddle point of $P_\mu^{*}$ and it maximizes $P_\mu^{*}$ in the direction $p_0^*$ such that $P_\mu^{*}(\bar{p}^*) = \sup_{\theta > 0} P_\mu^{*}(\bar{p}^* + \theta p_0^*)$, then substituting $p^* = \bar{p}^* + \theta p_0^*$ into (61) we also get a contradiction. Thus, the critical point $\bar{p}^*$ of $P_\mu^{*}$ must be a local minimizer on $E^*_b$.

Now if $(\bar{u}, \bar{p}^*)$ mini-maximizes $L_\mu$ on $U_0 \times E^*_b$ in the order of $\inf_{p^*} \sup_{a}$, then

$$ L_\mu(\bar{u}, \bar{p}^*) = \inf_{p^*} \sup_{a} L_{\mu}(u, p^*) = \inf_{p^*} P_\mu^{*}(p^*). $$

Lemma 1 then shows that $P_\mu^{*}(\bar{p}^*) = L_\mu(\bar{u}, \bar{p}^*) = \inf P_\mu^{*}(p^*)$. Similarly, we can prove that the critical point $\bar{u}$ of $P_\mu$ must be a local minimizer on $U_0$.

Conversely, if $P_\mu(\bar{u}) = \inf P_\mu(u)$ and $P_\mu^{*}(\bar{p}^*) = \inf P_\mu^{*}(p^*)$ then, on $U_0 \times E^*_b$,

$$ P_\mu(\bar{u}) = \inf_{u \in U_0} P_\mu(u) = \inf_{u} \sup_{p^*} L_{\mu}(u, p^*), $$

$$ P_\mu^{*}(\bar{p}^*) = \inf_{p^* \in E^*_b} P_\mu^{*}(p^*) = \inf_{p^* \in E^*_b} \sup_{u} L_{\mu}(u, p^*). $$

By Lemma 1, the condition $P_\mu(\bar{u}) = P_\mu^{*}(\bar{p}^*)$ shows that the critical point $(\bar{u}, \bar{p}^*)$ of $L_\mu$ mini-maximizes $L_\mu$ in either order on $U_0 \times E^*_b$. If $W : E \to \overline{R}$ is concave, the proof is similar to that of Theorem 6.

In non-convex systems, the Lagrangian $L_\mu$ could have several critical points. The gap function could be positive at one critical point and negative at others. Combining the above results, we have following interesting result (Gao 1997).

**Theorem 8**  (Triality theorem I)  Suppose that $(\bar{u}, \bar{p}^*) \in U_0 \times E^*_b$ is a critical point of $L_\mu$. If $W : E \to \overline{R}$ is convex and $G(\bar{u}, \bar{p}^*) \geq 0$ then

$$ P_\mu(\bar{u}) = \inf_{u \in U_0} P_\mu(u) \quad \text{if and only if} \quad P_\mu^{*}(\bar{p}^*) = \sup_{p^* \in E^*_b} P_\mu^{*}(p^*). \quad (62) $$
If $W : \mathcal{E} \to \overline{\mathbb{R}}$ is concave and $G(\bar{u}, \bar{p}^*) \leq 0$ then
\begin{equation}
 P_\mu(\bar{u}) = \sup_{u \in \mathcal{U}_b} P_\mu(u) \quad \text{if and only if} \quad P_\mu^*(\bar{p}^*) = \inf_{p^* \in \mathcal{E}_b^*} P_\mu^*(p^*). \tag{63}
\end{equation}

If $W : \mathcal{E} \to \overline{\mathbb{R}}$ is convex and $G(\bar{u}, \bar{p}^*) < 0$ or $W : \mathcal{E} \to \overline{\mathbb{R}}$ is concave and $G(\bar{u}, \bar{p}^*) > 0$ then, either
\begin{equation}
 P_\mu(\bar{u}) = \inf_{u \in \mathcal{U}_b} P_\mu(u) \quad \text{if and only if} \quad P_\mu^*(\bar{p}^*) = \inf_{p^* \in \mathcal{E}_b^*} P_\mu^*(p^*), \tag{64}
\end{equation}
or
\begin{equation}
 P_\mu(\bar{u}) = \sup_{u \in \mathcal{U}_b} P_\mu(u) \quad \text{if and only if} \quad P_\mu^*(\bar{p}^*) = \sup_{p^* \in \mathcal{E}_b^*} P_\mu^*(p^*). \tag{65}
\end{equation}

**Proof.** We give the proof only for convex $W$. Since $\Lambda : \mathcal{U} \to \mathcal{E}$ is quadratic, by Theorem 3 we know that the gap function $G(\bar{u}, \bar{p}^*) \geq 0$ if and only if $(\bar{u}, \bar{p}^*)$ is a saddle point of $L_\mu$ on $\mathcal{U}_b \times \mathcal{E}_b^*$. So, Theorem 5 shows that $\bar{u}$ minimizes $P_\mu$ on $\mathcal{U}_b \subset \mathcal{U}_d$ if and only if $\bar{p}^*$ maximizes $P_\mu^*$ on $\mathcal{E}_b^* \subset \mathcal{E}_u^*$.

If the gap function $G(\bar{u}, \bar{p}^*) < 0$, $(\bar{u}, \bar{p}^*)$ is a super-critical point of $L_\mu$ on $\mathcal{U}_b \times \mathcal{E}_b^*$, that is
\begin{equation}
 L_\mu(\bar{u}, p^*) \leq L_\mu(\bar{u}, \bar{p}^*) \geq L_\mu(u, \bar{p}^*) \quad \forall (u, p^*) \in \mathcal{U}_b \times \mathcal{E}_b^*.
\end{equation}
If $\bar{u}$ maximizes $P_\mu$ on $\mathcal{U}_b$, we have
\begin{align*}
P_\mu(\bar{u}) &= \sup_{u \in \mathcal{U}_b} P_\mu(u) = \sup_{u \in \mathcal{U}_b} \sup_{p^*} L_\mu(u, p^*) \\ &= \sup_{p^*} \sup_{u \in \mathcal{U}_b} L_\mu(u, p^*) = \sup_{p^*} P_\mu^*(p^*) = P_\mu^*(\bar{p}^*).
\end{align*}
This shows that $P_\mu(\bar{u}) = \sup_{u \in \mathcal{U}_b} P_\mu(u)$ is equivalent to $P_\mu^*(\bar{p}^*) = P_\mu^*(p^*)$.

If $\bar{u}$ minimizes $P_\mu$ on $\mathcal{U}_b$,
\begin{equation}
P_\mu(\bar{u}) = \inf_{u \in \mathcal{U}_b} P_\mu(u) = \inf_{u \in \mathcal{U}_b} \sup_{p^* \in \mathcal{E}_b^*} L_\mu(u, p^*) = L_\mu(\bar{u}, \bar{p}^*),
\end{equation}
that is, $(\bar{u}, \bar{p}^*)$ mini-maximizes $L_\mu$ on $\mathcal{U}_b \times \mathcal{E}_b^*$. By Theorem 7, $p^*$ minimizes $P_\mu^*$ on $\mathcal{E}_b^*$. Conversely, if $p^*$ minimizes $P_\mu^*$ on $\mathcal{E}_b^*$ then $(\bar{u}, \bar{p}^*)$ mini-maximizes $L_\mu$ on $\mathcal{U}_b \times \mathcal{E}_b^*$.

Theorem 9 (Triality theorem II) (a) Suppose that $W : \mathcal{E} \to \overline{\mathbb{R}}$ is convex and $(\bar{u}, \bar{p}^*) \in \mathcal{U}_b \times \mathcal{E}_b^*$ is a critical point of $L_\mu$. If $G(\bar{u}, \bar{p}^*) > 0$ then, on $\mathcal{U}_b \times \mathcal{E}_b^*$,
\begin{equation}
 \inf_{u \in \mathcal{U}_b} \sup_{p^*} L_\mu(u, p^*) = L_\mu(\bar{u}, \bar{p}^*) = \sup_{p^* \in \mathcal{E}_b^*} \inf_{u \in \mathcal{U}_b} L_\mu(u, p^*). \tag{66}
\end{equation}
If $G(\bar{u}, \bar{p}^*) < 0$ then, on $\mathcal{U}_b \times \mathcal{E}_b^*$, either
\begin{equation}
 \inf_{u \in \mathcal{U}_b} \sup_{p^*} L_\mu(u, p^*) = L_\mu(\bar{u}, \bar{p}^*) = \inf_{p^* \in \mathcal{E}_b^*} \sup_{u \in \mathcal{U}_b} L_\mu(u, p^*), \tag{67}
\end{equation}
or
\[ \sup_u \sup_{p^*} L_\mu(u, p^*) = L(\tilde{u}, \tilde{p}^*) = \sup_u \sup_{p^*} L_\mu(u, p^*). \] (68)

(b) Suppose that \( W : \mathcal{E} \to \mathbb{R} \) is concave and \( (\tilde{u}, \tilde{p}^*) \in \mathcal{U}_0 \times \mathcal{E}_0^* \) is a critical point of \( L_\mu \).

If \( G(\tilde{u}, \tilde{p}^*) \leq 0 \) then, on \( \mathcal{U}_0 \times \mathcal{E}_0^* \),
\[ \sup_u \inf_{p^*} L_\mu(u, p^*) = L_\mu(\tilde{u}, \tilde{p}^*) = \inf_u \sup_{p^*} L_\mu(u, p^*). \] (69)

If \( G(\tilde{u}, \tilde{p}^*) > 0 \) then, on \( \mathcal{U}_0 \times \mathcal{E}_0^* \), either
\[ \sup_u \inf_{p^*} L_\mu(u, p^*) = L_\mu(\tilde{u}, \tilde{p}^*) = \sup_u \inf_{p^*} L_\mu(u, p^*), \] (70)

or
\[ \inf_u \sup_{p^*} L_\mu(u, p^*) = L(\tilde{u}, \tilde{p}^*) = \inf_u \sup_{p^*} L_\mu(u, p^*). \] (71)

Proof. This theorem follows by combining Lemma 1 and Triality theorem I.

\[ \square \]

5. Applications in finite deformation theory

Let \( \Omega \subset \mathbb{R}^n \) be an open, simply connected, bounded domain with boundary \( \partial\Omega = \Gamma_I \cup \Gamma_u \), \( \Gamma_I \cap \Gamma_u = \emptyset \). The finite deformation from \( \Omega \) into \( \mathbb{R}^m \) is a mapping: \( \phi : \Omega \to \mathbb{R}^m \). The configuration variable space \( \mathcal{E} \) is a finite deformation vector space \( \mathcal{E} = \{ \phi \in L^\beta(\Omega; \mathbb{R}^m) \} \).

Its dual space \( \mathcal{E}^* = \{ f \in L^\beta(\Omega; \mathbb{R}^m) \} \) is a force space. Here \( L^\alpha(\Omega; \mathbb{R}^m) \) is the standard Lebesgue integrable space with domain \( \Omega \) and range \( \mathbb{R}^m \), \( \alpha, \beta \in [1, +\infty] \) are dual numbers: \( 1/\alpha + 1/\beta = 1 \). If \( \mathbf{f} \in \mathcal{E}^* \) is specified as the body force \( \mathbf{b} \) in \( \Omega \), and boundary force \( \mathbf{t} \) on \( \partial\Omega \), then the bilinear form \( \langle \phi, \mathbf{f} \rangle : \mathcal{U} \times \mathcal{E}^* \to \mathbb{R} \) represents the external work:

\[ \langle \phi, \mathbf{f} \rangle = \int_{\Omega} \phi \cdot \mathbf{b} \, d\Omega + \int_{\partial\Omega} \phi \cdot \mathbf{t} \, d\Gamma. \]

For mixed boundary value problems, we assume that on \( \Gamma_I \), the surface traction \( \mathbf{t} \) is given, while on the remaining part \( \Gamma_u \), the deformation \( \phi \) is prescribed. Then, all \textit{kinetically admissible deformations} \( \phi \in \mathcal{U} \) form a convex set \( \mathcal{C} \). For example, we can simply let

\[ \mathcal{C} = \{ \phi \in L^\alpha(\Omega; \mathbb{R}^m) \mid \phi(x) = \phi(x) \, \forall x \in \Gamma_u \}. \] (72)

The indicator function \( \psi_C \) of a set \( C \) is defined by

\[ \psi_C(\phi) = \begin{cases} 0 & \text{if } \phi \in C, \\ +\infty & \text{otherwise}, \end{cases} \]

which is convex, l.s.c. if \( C \) is convex and closed. Then, on the whole space \( \mathcal{U} \), the external energy \( F \) can be written as

\[ F(\phi) = \langle \phi, \mathbf{f} \rangle - \psi_C(\phi) = \int_{\Omega} \phi \cdot \mathbf{b} d\Omega + \int_{\Gamma_I} \phi \cdot \mathbf{t} d\Gamma - \psi_C(\phi). \] (74)
For a given geometric operator \( \Lambda \), we let \( V : \mathcal{D} \subset \mathcal{E} \to \mathbb{R} \) be the so-called stored energy density such that on the feasible space \( \mathcal{U}_a = \{ \phi \in \mathcal{U} \mid \phi \in \mathcal{C}, \ \Lambda \phi \in \mathcal{D} \} \), the primal problem is

\[
(P_{\text{int}}) : \quad P_\mu(\phi) = W_\mu(\Lambda \phi) - F(\phi) = \int_\Omega V(\Lambda \phi - \mu)\,d\Omega - \int_\Omega \phi \cdot \mathbf{b} \,d\Omega - \int_{\Gamma_f} \phi \cdot \mathbf{i} \,d\Gamma
\]

\[
\rightarrow \min \quad \forall \phi \in \mathcal{U}_a.
\] (75)

First, we let \( p = \Lambda \phi = \nabla \phi \). Then \( \Lambda : \mathcal{U} \to \mathcal{E} \) is a geometrically linear operator. In finite deformation theory, \( p = \nabla \phi \in \mathbb{R}^{m \times n} \) is called the deformation gradient, which is a two-point tensor, denoted by \( F \). We use Lin to denote the linear space of all second-order tensors. For admissible deformation, we define

\[
\mathcal{D}_F = \{ F \in \text{Lin}(\Omega; \mathbb{R}^{m \times n}) \mid \text{rank} F = \min\{m, n\}, \quad V(F) \in L^1(\Omega; \mathbb{R}) \}. \quad (76)
\]

The conjugate variable of \( F \) is the so-called first Piola–Kirchhoff stress tensor defined by

\[
\mathbf{\tau} = D\mathbf{V}(\mathbf{F}),
\]

which is also a two-point tensor. Since \( \mathbf{F} \) is not a strain measure, \( V(\mathbf{F}) \) is in general non-convex (see, for example, Ogden, 1984), so the duality relationship between \( \mathbf{\tau} \) and \( \mathbf{F} \) is not one-to-one, and \((P_{\text{int}})\) cannot be considered as a geometrically linear problem. The parameter \( \mu \in \mathbb{R}^{m \times n} \) in this problem can be considered as a stress free deformation state or as an internal variable of the material. The critical-point condition \( DP_\mu(\phi) = 0 \) leads to the following mixed boundary value problem.

**PROBLEM 4** For a given \( \mu \in \text{Lin}(\Omega; \mathbb{R}^{m \times n}) \), find \( \phi \in \mathcal{U}_a \) and the associated \( \mathbf{\tilde{\tau}} = D\mathbf{V}^*(\nabla \phi - \mu) \) such that

\[
(BVP) : \quad \Lambda^* \mathbf{\tilde{\tau}}(\phi) = D\mathbf{V}(\phi) \Rightarrow \begin{cases}
-\nabla \cdot (D\mathbf{V}^*(\nabla \phi)) = \mathbf{b} & \text{in } \Omega, \\
\mathbf{n} \cdot (D\mathbf{V}^*(\nabla \phi)) = \mathbf{i} & \text{on } \Gamma_i,
\end{cases}
\] (77)

where \( \mathbf{n} \) is the unit vector of the external normal to \( \Gamma_i \).

By the Fenchel–Rockafellar duality, we have

\[
W^*_\mu(\mathbf{\tau}) = \sup_{\mathbf{F} \in \mathcal{D}} \left\{ \int_\Omega \text{tr}(\mathbf{\tau}^T \cdot \mathbf{F})\,d\Omega - \int_\Omega V(\mathbf{F} - \mu)\,d\Omega \right\}
\]

\[
= \int_\Omega \left[ V^*(\mathbf{\tau}) + \text{tr}(\mathbf{\tau}^T \cdot \mu) \right]\,d\Omega + \psi_{D^*}\mathbf{\tau},
\] (78)

\[
F^*(\Lambda^* \mathbf{\tau}) = \inf_{\phi \in \mathcal{C}} \left\{ (\phi, \Lambda^* \mathbf{\tau}) - \int_\Omega \phi \cdot \mathbf{b} \,d\Omega - \int_{\Gamma_f} \phi \cdot \mathbf{i} \,d\Gamma \right\}
\]

\[
= \int_{\Gamma_u} \mathbf{\tilde{\phi}} \cdot \mathbf{\tau} \cdot \mathbf{n} \,d\Gamma - \psi_{C^*}(\Lambda^* \mathbf{\tau}),
\] (79)

where \( D^*_\mathbf{\tau} \) and \( C^*_\mathbf{\tau} \) are defined by

\[
D^*_\mathbf{\tau} = \{ \mathbf{\tau} \in DV(D_F) \mid V^*(\mathbf{\tau}) \in L^1(\Omega; \mathbb{R}) \},
\]

\[
C^*_\mathbf{\tau} = \{ f \in L^0(\Omega; \mathbb{R}^m) \mid f(x) = \mathbf{b}(x) \forall x \in \Omega, \quad f(x) = \mathbf{i} \quad \forall x \in \Gamma_i \}.
\]
The traditional Lagrangian form of \((\mathcal{P}_{\inf})\) is
\[
L_{L-Z}(\phi, \tau) = \int_{\Omega} \left[ \text{tr}(\tau^T \cdot (\nabla \phi)) - V_\mu^*(\tau) - \bar{b} \cdot \phi \right] d\Omega - \int_{\Gamma_n} \tilde{t} \cdot \phi d\Gamma. \tag{81}
\]

Unfortunately, the critical condition \(DL_{L-Z}(\phi, \tau) = 0\) is not equivalent to \((\text{BVP})\) and \(P_\mu(\phi) \geq \sup_\tau L_{L-Z}(\phi, \tau)\) because of the non-convexity of \(V(F)\).

Let \(T_a = \mathcal{E}_a^*\) be the so-called statically admissible space defined by
\[
T_a = \{ \tau \in \mathcal{D}_a^* \mid -\nabla \cdot \tau = \bar{b} \quad \text{in} \quad \Omega; \quad \mathbf{n} \cdot \tau^T = \tilde{t} \quad \text{on} \quad \Gamma_t \}. \tag{82}
\]

On \(C \times T_a\), \(L_{L-Z}\) reduces to the Fenchel–Rockafellar conjugate functional for the geometrically linear operator \(A\phi = \nabla \phi\). Thus,
\[
P_{L-Z}^c(\tau) = \int_{\Gamma_n} \phi \cdot \tau \cdot \mathbf{n} d\Gamma - \int_\Omega V_\mu^*(\tau) d\Omega. \tag{83}
\]

This is the well-known Levinson–Zubov energy in finite deformation theory. It is obvious that \(P_{L-Z}^c : T_a \to \mathbb{R}\) is always concave and upper semicontinuous. So, we have the Levinson–Zubov complementary variational problem
\[
(\mathcal{P}_{L-Z}^c) : \quad P_{L-Z}^c(\tau) = \sup_{\phi \in \mathcal{D}_a} P_{L-Z}^c(\tau). \tag{84}
\]

This problem is not equivalent to the primal problem and there exists a duality gap between \((\mathcal{P}_{\inf})\) and \((\mathcal{P}_{L-Z}^c)\) (see Gao 1992, 1989b), that is,
\[
\inf_{\phi \in \mathcal{D}_a} P(\phi) \geq \sup_{\tau \in T_a} P_{L-Z}^c(\tau). \tag{85}
\]

The equality holds if and only if \(V(F)\) is convex.

Now we let \(A\) be a quadratic operator such that \(A\phi\) is the so-called right Cauchy–Green strain tensor
\[
C = A\phi = \frac{1}{2}F^T \cdot F = \frac{1}{2}(\nabla \phi)^T \cdot (\nabla \phi) \in \mathbb{R}^{n \times n}. \tag{86}
\]

Then \(A_I\) and \(A_n\) can be given as
\[
A_I(\phi)\phi = \frac{1}{2}[(\nabla \phi)^T \cdot (\nabla \phi) + (\nabla \phi)^T \cdot (\nabla \phi)],
\]
\[
A_n(\phi)\phi = -\frac{1}{2}[(\nabla \phi)^T \cdot (\nabla \phi) + (\nabla \phi)^T \cdot (\nabla \phi)].
\]

Let
\[
\mathcal{D} := \{ C \in \text{Lin}(\Omega; \mathbb{R}^{n \times n}) \mid C = C^T, \quad \det C > 0, \quad V(C) \in L^1(\Omega; \mathbb{R}) \text{ is G-differentiable} \}.
\]

For hyperelasticity, we assume that the strain energy density \(V(C)\) is convex. The conjugate variable of \(C\) is the second Piola–Kirchhoff stress \(p^* = T = D V(C)\). Denoting the parameter \(\mu = C_0 \in \mathbb{R}^{n \times n}\), we have the complementary energy
\[
W_{\mu}^*(T) = \sup_{C \in \mathcal{D}} \{ \langle C, T \rangle - \int_\Omega V(C - C_0) d\Omega \}
\]
\[
= \int_\Omega [V^*(T) + \text{tr}(T \cdot C_0)] d\Omega + \psi_{\mathcal{D}}^*(T),
\]

which is convex and l.s.c., where

\[ D^* := \{ T \in DV(D) \mid T = T^T, \quad \det T \neq 0, \quad V^*(T) \in L^1(\Omega; \mathbb{R}) \text{ is G-differentiable} \}. \]

On \( D \times D^* \), we have the following relations:

\[ T = DV(C) \iff C = DV^*(T) \iff V(C) + V^*(T) = \text{tr}(C \cdot T). \quad (87) \]

The critical condition \( DP_{\mu}(\bar{\phi}) = 0 \) leads to the following boundary value problem.

**PROBLEM 5** For a given \( C_0 \in \text{Lin}(\Omega; \mathbb{R}^{n \times n}) \) and external force \( \bar{f} \in C^*_0 \), find \( \bar{\phi} \in U_0 \) and the associated \( \bar{T} = DV(C(\bar{\phi}) - C_0) \) such that

\[ (\text{BVP}) : \quad A^*_T(\bar{\phi})T = DF(\bar{\phi}) \Rightarrow \begin{cases} -\nabla \cdot (\nabla \bar{\phi} \cdot \bar{T}) = \bar{b} & \text{in } \Omega, \\ n \cdot (\nabla \bar{\phi} \cdot \bar{T}) = \bar{f} & \text{on } \Gamma_t. \end{cases} \quad (88) \]

Since \( \tau = F \cdot T \), the two boundary value problems (77) and (88) are equivalent. The gap function for the quadratic geometrical operator is a quadratic functional of \( \phi \), namely

\[ G(\phi, T) = (T, -A_T(\phi)\phi) = \frac{1}{2} \int_{\Omega} \text{tr}(\nabla \phi) \cdot T \cdot (\nabla \phi)^T \, d\Omega. \quad (89) \]

On \( C \times D^* \) the Lagrangian for the strain measure \( C(\phi) \) is

\[ L_{\mu}(\phi, T) = \int_{\Omega} [\text{tr}(C(\phi) \cdot T) - V^*_\mu(T)]d\Omega - \int_{\Omega} \phi \cdot \bar{b} \, d\Omega - \int_{\Gamma_t} \phi \cdot \bar{f} \, d\Gamma. \quad (90) \]

Since \( V(C) \) is convex, we have \( P_{\mu}(\phi) = \sup_T L_{\mu}(\phi, T) \).

By Theorem 4, the pure complementary energy for the finite deformation problem \( (\mathcal{P}_{\text{inf}}) \) is

\[ P^*_{\mu}(T) = \begin{cases} \inf_{\phi \in U_0} L_{\mu}(\phi, T) & \text{if } G(\phi, T) \geq 0, \\ \sup_{\phi \in U_0} L_{\mu}(\phi, T) & \text{if } G(\phi, T) < 0. \end{cases} \quad (91) \]

For any given \( \tau \in T_a \), the solution for (91) is \( \bar{\phi} = \tau \cdot T^{-1} \forall \ T \in D^* \). Thus, the pure complementary energy can be expressed as

\[ P^+_{\tau}(T) = \int_{\Gamma_a} \bar{\phi} \cdot \tau \cdot n \, d\Gamma - \int_{\Omega} V^*_\mu(T) \, d\Omega - G^*(T, \tau), \quad (92) \]

where \( G^*(T, \tau) = G(\phi(\tau, T), T) \) is given by

\[ G^*(T, \tau) = \int_{\Omega} \frac{1}{2} \text{tr}[\tau \cdot T^{-1} \cdot \tau^T] \, d\Omega. \quad (93) \]

The critical condition \( DP^*_{\tau}(\bar{T}) = 0 \) gives the dual Euler–Lagrange equation

\[ \bar{T} \cdot [DV^*(\bar{T}) + C_0] \cdot \bar{T} = \frac{1}{2} \tau^T \cdot \tau \quad \text{in } \Omega. \quad (94) \]

This is an algebraic equation! We therefore have the following result.
THEOREM 10 (Analytic solution theorem) Suppose that for a given $\phi(x) \forall x \in \mathcal{I}$ and the external load $f \in C^*_E$ such that the statically admissible space $\mathcal{T}_a$ is not empty. If for a given $\bar{t} \in \mathcal{T}_a$ such that $\bar{T} \in D^*$ is a solution of the dual Euler–Lagrange equation (94), and $\nabla \times (\bar{t} \cdot \bar{T}^{-1}) = 0$, then along any integral path from $x_0$ to $x$, the deformation given by

$$\phi(x) = \int_{x_0}^x \bar{t} \cdot \bar{T}^{-1} \cdot d\bar{x} + \phi(x_0)$$  \hspace{1cm} (95)$$

is a generalized solution of the (BVP). If $G(\phi, \bar{T}) \geq 0$, then $\phi(x)$ is a local minimizer of $P_{\mu}(\phi)$. The solution $\phi$ is a global minimizer of $P_{\mu}(\phi)$ if $G^*(\bar{T}, \tau) \geq 0 \ \forall \tau \in \mathcal{T}_a$.

Proof. With the Lagrange multiplier $\phi \in C$ introduced to relax the statically admissible conditions in $\mathcal{T}_a$, the Lagrangian $L_\tau : C \times D^* \times \text{Lin} \to \mathbb{R}$ associated with $P^*_\mu$ is

$$L_\tau(\phi, \bar{T}, \tau) = \int_{\Omega} [\text{tr}(\nabla \phi^T \cdot \tau) - V^*_\mu(\bar{T}) - \frac{1}{2}\text{tr}[\tau \cdot \bar{T}^{-1} \cdot \tau^T]]d\mathbb{Q} - \int_{\Omega} \phi \cdot d\mathbb{Q} - \int_{\Gamma_1} \phi \cdot d\mathbb{Q}.$$ \hspace{1cm} (96)

The critical-point condition $DL_\tau(\phi, \bar{T}, \tau) = 0$ gives the equations

$$\nabla \cdot \bar{t} + \bar{t} = 0 \quad \text{in} \ \Omega, \quad n \cdot \bar{t} = \bar{t} \quad \text{on} \ \Gamma_1,$$ \hspace{1cm} (97)

$$\nabla \phi = \bar{t} \cdot \bar{T}^{-1} \quad \text{in} \ \Omega,$$ \hspace{1cm} (98)

$$DV^*_\mu(\bar{T}) = \frac{1}{2}(\tau \cdot \bar{T}^{-1})^T \cdot (\bar{t} \cdot \bar{T}^{-1}) \quad \text{in} \ \Omega.$$ \hspace{1cm} (99)

For a given $\bar{t} \in \mathcal{T}_a$, substituting (98) into (99), we get

$$DV^*_\mu(\bar{T}) = \frac{1}{2}(\nabla \phi)^T \cdot (\nabla \phi) - C_0.$$ 

By the convexity of $V(C)$, this inverse constitutive equation is equivalent to $\bar{T} = DV(\bar{C}(\phi)) = C_0$. This shows that the critical point $(\phi, \bar{T}, \bar{t})$ of $L_\tau(\phi, \bar{T}, \tau)$ solves the (BVP). If $\nabla \times (\bar{t} \cdot \bar{T}^{-1}) = 0$, then $\nabla \phi$ defined by (98) is a conservative field. By vector analysis we know that the integration in (95) is path independent, and hence, $\phi$ is a solution of (BVP). By Lemma 1, substituting the extremality condition $\bar{t} = (\nabla \phi) \cdot \bar{T} \in \mathcal{T}_a$ into $P^*_\mu$ leads to $P_{\mu}(\phi) = L_\mu(\phi, \bar{T}) = P^*_\mu(\bar{T})$. By triality theorem II, if the gap function $G(\phi, \bar{T}) \geq 0$, then $\phi$ is a local minimizer. If $G^*(\bar{T}, \tau) \geq 0 \ \forall \tau \in \mathcal{T}_a$, then all critical points of $L_\mu$ should be saddle points. By Lemma 1 and Theorem 1, $\phi$ is a global minimizer of $P_{\mu}$. But $P_{\mu}$ is strictly convex if $G^*(\bar{T}, \tau) > 0 \ \forall \tau \in \mathcal{T}_a$. In this case, $\phi$ is a unique minimizer of $(P_{\mu})$. \hfill \square

According to Hill (1978), for a given $k \in \mathbb{R}$, the stored energy density $V$ should be a convex function of the Seth–Hill strain family

$$E^{(k)} := \frac{1}{2k}[(\bar{T}^k \cdot \bar{F}^k - I)].$$  \hspace{1cm} (100)$$

The conjugate stress tensor is then $T^{(k)} = DV(E^{(k)})$. In hyperelasticity the geometric operator $A^{(k)}$ is usually nonlinear, and hence the problem is fully nonlinear.
\( u(x) + x \); then \( u : \Omega \rightarrow \mathbb{R}^n \) is a displacement vector. The generalized Lagrangian for Seth–Hill strain family was given in Gao (1992) as

\[
L^{(k)}(u, T^{(k)}) = \int_{\Omega} \left[ \text{tr}(T^{(k)} \cdot E^{(k)}(u)) - W^*(T^{(k)}) - f \cdot u \right] d\Omega - \int_{\Gamma} t \cdot u d\Gamma. \tag{101}
\]

If \( k = 1 \), \( E^{(1)} = C - \frac{1}{2} I = \mathbf{E} \) is the well-known Green strain tensor. Its conjugate stress \( T^{(1)} \) is equal to \( \mathbf{T} \). The geometrical operator \( \Lambda \) in this case is the quadratic operator

\[
\mathbf{E} = \Lambda \mathbf{u} = \frac{1}{2} [\nabla \mathbf{u} + (\nabla \mathbf{u})^T + (\nabla \mathbf{u})^T \cdot (\nabla \mathbf{u})], \tag{102}
\]

and

\[
\Lambda_n(\bar{u}) \mathbf{u} = -\frac{1}{4} [\nabla \bar{u}]^T \cdot (\nabla \mathbf{u}) + (\nabla \mathbf{u})^T \cdot (\nabla \bar{u})].
\]

For the St-Venant–Kirchhoff material, the density of a stored energy \( V(\mathbf{E}) \) is the quadratic function (see, for example, (Ciarlet 1988))

\[
V(\mathbf{E}) = \frac{1}{2} \lambda (\text{tr} \mathbf{E})^2 + \nu \text{tr} \mathbf{E}^2, \tag{103}
\]

where \( \lambda, \nu \) are Lamé constants. Then, \( W(\Lambda \mathbf{u} - \mu) \) is a double-well energy of \( \nabla \mathbf{u} \). The gap function in this example is

\[
G(u, \mathbf{T}) = \int_{\Omega} \frac{1}{2} \text{tr} [\nabla \mathbf{u} \cdot (\nabla \mathbf{u})^T] d\Omega. \tag{104}
\]

On the equilibrium admissible space

\[
\mathcal{E}_u^* = \{ (\mathbf{u}, \mathbf{T}) \in \mathcal{C} \times \text{Lin} | \Lambda_n^*(\mathbf{u}) \mathbf{T} = DF(\mathbf{u}) \},
\]

the Lagrangian \( L^{(1)} \) can be written as

\[
P_H^{\mathcal{E}_u^*}(\mathbf{T}, \mathbf{u}) = \int_{\Gamma} \bar{u} \cdot (F(\mathbf{u}) \cdot \mathbf{T}) \cdot n d\Gamma - \int_{\Omega} V^*(\mathbf{T}) d\Omega - G(u, \mathbf{T}). \tag{105}
\]

This is the well-known Hellinger–Reissner complementary energy (Reissner, 1953), which has many consequences in mechanics (see, for example, Atluri 1980; Lee & Shield 1980; Nemat-Nasser 1972; Oden & Reddy 1983; Ogden 1977; Ogden 1984). Since \( P_H^{\mathcal{E}_u^*} \) depends on both \( \mathbf{u} \) and \( \mathbf{T} \), it is not considered as a pure complementary energy principle. The extremum property of the generalized Hellinger–Reissner energy \( L^{(1)}(\mathbf{u}, \mathbf{T}) \) was an open problem for many years. This problem is now solved by the triality theorems.

In a series of publications by Telega and co-workers, a method has been used repeatedly for obtaining what they called complementary energies in nonlinear mechanics (see Galka & Telega 1992). The key point of this method is to split the Green strain operator \( \Lambda \mathbf{u} = \mathbf{E} \) into a linear part \( \frac{1}{2} [\nabla \mathbf{u} + (\nabla \mathbf{u})^T] \) and a quadratic part \( \frac{1}{4} (\nabla \mathbf{u})^T \cdot (\nabla \mathbf{u}) \) (see Telega 1989). Unfortunately, all these very complicated complementary energies produced by this method are a trivial functional (identical to \( +\infty \)) unless the second Piola–Kirchhoff stress \( \mathbf{T} \) is positive definite almost everywhere in \( \Omega \). It is well known that if \( \mathbf{T}(\mathbf{x}) \) is positive definite in \( \Omega \),
the total potential $P(u)$ is convex (see Theorem 1) and the classical complementary energy principles are equivalent to the Levinson–Zubov principle, which has the simplest formulation (see, Gao 1992, Theorem 5). Moreover, a key mistake in this method was pointed out in Gao (1997). Applications of the Levinson–Zubov principle have been well studied in large deformation structures (see, for example, Stumpf 1979; Wempner 1986).

It was proved in Gao & Strang (1989a) that if the gap function $G(u, T)$ is positive on $E^*_u$, then $L^{(1)}$ is a saddle functional. By the triality theorems, we know that if this gap function is negative on $E^*_u$ then $L^{(1)}$ is a super critical-point functional. In this case, the problem has two dual extremum principles. Of course, different deformation operators $\Lambda$ lead to different complementary energy principles. The associated gap functions provide global extremum criteria for both primal and dual problems. A unified approach and generalized variational principles for the Seth–Hill strain operators were discussed in Gao (1992,1998a).

In differential geometry, $m = n + 1$, the finite deformation $\phi(x): \Omega \subset \mathbb{R}^n \to \mathbb{R}^{n+1}$ is a hyper-surface; $C \in \mathcal{D}$ is the well-known Riemannian metric tensor. If $\Phi = 0$, and $V(C) = (\det C)^{1/2}$, then the analytic solution $\phi$ defined by equation (95) is a minimum surface. In this case, the energy density $V(C)$ is concave. But in terms of $F = \nabla \phi$, $V(F) = (\det(F^T \cdot F))^{1/2}$ is convex. The multi-duality in minimal surface type problems was studied in Gao & Yang (1995). It is interesting to point out that the solution of the so-called complementary problem of $(P_{\inf})$ is the conjugate (or adjoint) surface of the minimal surface (see Gao 1999a).

6. One-dimensional non-convex variational problems

Using the results presented in this paper, we are able to solve the one-dimensional non-convex variational problem

$$(P_m) : \quad P_m(u) = \int_0^1 \frac{1}{m} E(\Lambda u - \mu)^m \, dx - \int_0^1 f(x)u \, dx \to \min \forall u \in \mathcal{U}_a, \quad (106)$$

Let $\mathcal{D} = L^a(0, 1)$ with $a \in [1, \infty]$. The stored energy is

$$W_m(p) = \int_0^1 \frac{1}{m} E(p - \mu)^m.$$

Let $p^* = \sigma$, so that the duality relation can be written as

$$\sigma = DW_m(p) = E(p - \mu)^{m-1}. \quad (107)$$

The external energy $F(u) = \int_0^1 f(x)u \, dx - \Psi_C(u)$ is a linear functional on a given convex set $C \subset L^2(0, 1)$. The convexity of the problem depends on the constant $m \in \mathbb{R}$ and the operator $\Lambda$. We consider the following three cases.

6.1 Unilateral bifurcation problem: $m = 2$, $\Lambda u = \frac{1}{2} (u_{,xx})^2 + c(x)u_{,x}$

For any given continuous function $c(x)$, $\Lambda$ is a quadratic operator and

$$\Lambda_1(u)u = \frac{1}{2} u_{,xx}^2 + cu_{,x}, \quad \Lambda_2(u)u = -\frac{1}{2} u_{,x}^2.$$
Since \( A \) is nonlinear, the problem \((P_m)\) is non-convex with a double-well energy

\[
V(p(\epsilon)) = \frac{1}{2} E(\frac{1}{2} \epsilon^2 + c \epsilon - \mu)^2.
\]

But in terms of \( p = Au \), the stored energy

\[
W(p - \mu) = \int_0^1 \frac{1}{2} E(p - \mu)^2 dx
\]

is a quadratic function and the constitutive equation \( \sigma = DW_\mu(p) = E(p - \mu) \) is linear.

For the unilateral variational problem, we let \( D_1 = L^2(0, 1) \) and

\[
C_1 = \{ u \in L^2(0, 1) \mid u(x) \geq 0 \ 0 \leq x \leq 0, \ 0 < 0 \}
\]

which is a convex cone. Then on the feasible space \( U_u = \{ u \in C_1 \mid Au \in D_1 \} \), the primal problem \((P_m)\) is equivalent to the following nonlinear complementarity problem (NCP):

\[
\begin{align}
\quad & u(x) \geq 0 \quad \forall x \in [0, 1], \quad (108) \\
\quad & [E(\frac{1}{2} u^2_{xx} + cu_x - \mu)(u_x + c)]_x + f \leq 0 \quad \forall x \in [0, 1], \quad (109) \\
\quad & u(x)[E(\frac{1}{2} u^2_{xx} + cu_x - \mu)(u_x + c)]_x + f = 0 \quad \forall x \in [0, 1]. \quad (110)
\end{align}
\]

The direct methods for solving this geometrically nonlinear complementarity problem are very difficult. But by using the duality theory, we are able to find the analytic solution.

The complementary energy in this problem is

\[
W_\mu^*(\sigma) = \sup_{p \in D_1} \left\{ \int_0^1 \sigma p dx - V(p - \mu) \right\}
\]

\[
= \left\{ \begin{array}{ll}
\int_0^1 \left( \frac{1}{2} E \sigma^2 + \mu \sigma \right) dx & \text{if } \sigma \in DW_\mu(D_1), \ \sigma \neq 0, \\
0 & \text{if } \sigma \in DW_\mu(D_1), \ \sigma = 0,
\end{array} \right.
\]

otherwise.

Note that \( \sigma = 0 \) implies that \( p = \mu \). If \( c = 0, \ p = \frac{1}{2} u^2_{xx} \geq 0 \ \forall u \in C, \) the range of \( DW_\mu(D_1) \) is \(-\mu E \leq \sigma < +\infty\). Then

\[
D_1^* = \{ \sigma \in L^2[0, 1] \mid -\mu E \leq \sigma < +\infty, \ \sigma(x) \neq 0 \ \forall x \in (0, 1) \}. \quad (111)
\]

For any given source function \( f(x) \), the conjugate function of \( F(u) \) should be

\[
F^*(u^*) = \inf_{u \in C_1^*} \left\{ \int_0^1 uu^* dx + uu^*|_{x=0} - \int_0^1 f u dx \right\}
\]

\[
= \left\{ \begin{array}{ll}
0 & \text{if } u^* \in C_1^*, \\
-\infty & \text{otherwise},
\end{array} \right.
\]

where

\[
C_1^* = \{ u^* \in L^2[0, 1] \mid u^*(x) \geq f(x) \ \forall x \in (0, 1), \ u^*(1) = 0 \}
\]

is a dual convex cone. Since \( W(p(\epsilon)) \) is non-convex in the linear deformation \( \epsilon = u_x \), the
Piola stress \( \tau \) is given by \( \tau = D_x W_\mu(p(\epsilon)) = \epsilon \sigma \). So, the statically admissible space \( \mathcal{T}_a \) in this unilateral variational problem is
\[
\mathcal{T}_a^+ = \{ \tau \in \mathcal{W}^{1,2}(0, 1) | -\tau, x \geq f(x), \forall x \in (0, 1), \tau(1) = 0 \}. \tag{112}
\]
It is easy to find that
\[
\tau(x) \leq \int_0^x f(t) \, dt + \int_0^1 f(t) \, dt, \quad \forall x \in [0, 1] \tag{113}
\]
is statically feasible.

The Lagrangian \( L_\mu : C^1 \times \mathcal{D}_1^* \to \mathbb{R} \) for this problem is
\[
L_1(u, \sigma) = \int_0^1 \left[ \frac{1}{2} u_x^2 \sigma - \frac{1}{2E} \sigma^2 - \mu \sigma - fu \right] \, dx. \tag{114}
\]
The gap function in this problem is a quadratic function of \( u \):
\[
G(u, \sigma) = \langle \sigma, -\Lambda u \rangle = \int_0^1 \frac{1}{2} \sigma u_x^2 \, dx.
\]
The pure complementary energy \( P^*_\mu \) for this one-dimensional problem is
\[
P^*_\mu(\sigma) = -\int_0^1 \left[ \frac{1}{2E} \sigma^2 + \mu \sigma + \frac{1}{2} \sigma^{-1} \tau^2 \right] \, dx, \tag{115}
\]
which is well defined on \( \mathcal{D}_1^* \). The dual Euler–Lagrange equation (94) in this example is the cubic algebraic equation
\[
\sigma^2 \left( \frac{1}{E} \sigma + \mu \right) = \frac{1}{2} \tau^2. \tag{116}
\]
In algebraic geometry, this equation is the so-called singular cubic curve (see Fig. 5.1 (b)), which is a special case of the well-known Weierstrass equation \( y^2 = x^3 + ax^2 + bx + c \).

All the non-singular points on this curve form an Abelian group (see Silverman & Tate 1992). For a given \( f(x) \) such that \( \tau(x) \) is obtained by (125), equation (116) has at most three solutions \( \sigma_i \) \( (i = 1, 2, 3) \). Since \( u(0) = 0 \), the analytic solution (95) in this example is then
\[
u_i(x) = \int_0^x \frac{\tau(t)}{\sigma_i(t)} \, dt. \tag{117}
\]
But only those \( u_i(x) \geq 0, \forall x \in [0, 1] \) solve the unilateral variational problem.

Let \( E = 1 \) and \( h(\sigma) = \sigma^3 + \mu \sigma^2 \). If \( \mu \leq 0 \), the gap function is positive on \( \mathcal{D}_1^* \).

In this case, \( P_\mu(u) \) is convex, and the algebraic equation (116) has only one real solution \( \sigma_1(x) \geq 0, \forall x \in [0, 1] \) (see Fig. 2 (b)). If \( \mu > 0 \), the gap function could be negative on \( \mathcal{D}_1^* \).

In this case, \( P_\mu(u) \) is non-convex. The cubic \( h(\sigma) \) has a local maximum \( h_{\max} = 4\mu^3 / 27 \) at \( \bar{\sigma} = -2\mu / 3 \), and \( h(\sigma) \leq h_{\max} \forall \sigma \leq 0 \) (see Fig. 2 (a)). For any given load \( f(x) \), \( \tau(x) \) is a continuous function on the closed interval \( [0, 1] \). Suppose that \( \tau_{\min}^2 \leq \tau^2(x) \leq \tau_{\max}^2 \), \( x \in [0, 1] \), then the extremality conditions \( h_{\max} = 1 / 2 \tau_{\min}^2 \) and \( h_{\max} = 1 / 2 \tau_{\max}^2 \) give
\[
\mu_{\min} = 1.5 \tau_{\min}^{2/3}, \quad \mu_{\max} = 1.5 \tau_{\max}^{2/3}. \tag{118}
\]
So, if $\mu \in (-\infty, \mu_{\text{min}})$, the algebraic equation has only one real solution $\sigma_1 > 0$. However, if $\mu \in (\mu_{\text{min}}, \mu_{\text{max}})$ there exists at least one point $x_0 \in (0, 1)$ such that the equation has three real roots $\sigma_i (i = 1, 2, 3)$ at this point, and one of them must be positive, while the other two are negative (see Fig. 3 (b)). For $x \in (x_0, 1]$, the equation may have two roots. If $\mu > \mu_{\text{max}}$, the algebraic equation has three real roots everywhere in $(0, 1)$, which give three critical points $u_i (i = 1, 2, 3)$ of the total potential $P_{\mu}(u)$. If $\sigma_1 > 0 > \sigma_2 > \sigma_3$, then by the trinity theory, $u_1$ is a global minimizer, $u_2$ is a local minimizer and $u_3$ is a local maximizer (see Fig. 5.1 (a)). For $f = \pm 1$, solutions $\sigma_i$ are shown in Fig. 3. If $f > 0$, the solution of this unilateral variational problem should be $u_1(x) > 0$, which is a global
minimizer (see Figure 5.1 (a)). However, if \( f < 0 \), the solution should be \( u_2(x) > 0 \), which is a local minimizer.

6.2 Concave stored energy: \( m = \frac{1}{2}, \quad p = \lambda u = u^2_{xx} \)

In this case, we let \( \mu = -\lambda < 0 \). Then, the primal problem

\[
P_\lambda(u) = \int_0^1 2\sqrt{u_{xx}^2 + \lambda} \, dx - \int_0^1 fu \, dx \rightarrow \min
\]

(119)
is a well-known problem in the calculus of variations and differential geometry. Since \( P_\lambda(u) \) is convex, its Fenchel–Rockafellar dual problem has been well studied for the geometrically linear operator \( \epsilon(u) = u_{xx} \) (see Ekeland & Temam 1976). Here we choose the quadratic geometric operator, the stored energy density \( V_\lambda(p) = 2(p + \lambda)^{1/2} \) to be concave. But our method can solve this problem in a very easy way. Let

\[
C_2 = \{ u \in L^2(0, 1) \mid u(0) = 0 \}, \quad D_2 = \{ p \in L^1(0, 1) \mid p(x) \geq 0 \ \forall x \in (0, 1) \}.
\]

The dual variable for \( p = (u_{xx})^2 \) is \( \sigma = DV_\lambda(p) = 1/(p + \lambda)^{1/2} \geq 0 \). The complementary energy \( W_\lambda^* \) is then

\[
W_\lambda^*(\sigma) = \inf_{p \in D_2} \left\{ \int_0^1 \left[ p\sigma - 2(p + \lambda)^{1/2} \right] \, dx \right\} = \int_0^1 \left[ -\lambda\sigma - \frac{1}{\sigma} \right] \, dx + \Psi_{D_2^*}(\sigma),
\]

(120)

where the feasible set \( D_2^* \) is

\[
D_2^* = \{ \sigma \in DV_\lambda(D_2) \mid 0 < \sigma \leq \sqrt{1/\lambda} \}.
\]

(121)

Then, on \( C \times D_2^* \), the Lagrangian is

\[
L_\lambda(u, \sigma) = \int_0^1 \left[ \sigma u_{xx}^2 + \lambda\sigma + \frac{1}{\sigma} - fu \right] \, dx.
\]

(122)

The critical condition \( DuL_\lambda = 0 \) gives the pure complementary energy

\[
P_\lambda^*(\sigma) = \int_0^1 \left[ \lambda\sigma + \frac{1}{\sigma} - \frac{\tau^2}{4\sigma} \right] \, dx.
\]

(123)

In this example, the statically admissible space \( T_\lambda \), given by

\[
T_\lambda = \{ \tau \in W^{1,2}(0, 1) \mid \tau_x + f(x) = 0 \ \forall x \in (0, 1), \ \tau(1) = 0 \}
\]

(124)

has a unique element, namely

\[
\tau(x) = \int_0^x -f(t) \, dt + \int_0^1 f(t) \, dt \ \forall x \in [0, 1].
\]

(125)

For a given \( f \), the dual Euler–Lagrange equation (94) has only the two roots

\[
\sigma = \pm \frac{1}{\sqrt{\lambda}} \sqrt{1 - \tau^2/4},
\]

(126)
On $D_2^*$, $P_h^*$ has only one critical point $\bar{\sigma} = (1/\sqrt{\lambda})(1 - \tau^2/4)^{1/2}$, which gives the unique solution:

$$\bar{u}(x) = \int_0^x \frac{\sqrt{\lambda \tau(t)}}{\sqrt{4 - \tau^2(t)}} \, dt.$$  \hspace{1cm} (127)

On $D_2^*$ the gap function $G(u, \sigma) = \int_0^1 \sigma u^2_x \, dx > 0 \ \forall u \neq 0$, and for any given load $f$ such that $|\tau| < 2$, $\bar{u}$ is a unique global minimizer of $P_h$ and $\inf P_h(u) = \inf P_h^*(\sigma)$.

The primal energy corresponding to $U^*(\sigma) = \lambda \sigma + 1/\sigma - \tau^2/4 \sigma$ is

$$U(u) = 2\sqrt{u^2 + \lambda} - \tau u.$$  \hspace{1cm} (128)

For a given $\lambda > 0$ and $\tau$, $U$ is convex and the equality $\inf U = \inf U^*$ happens only for $\sigma = (1/\sqrt{\lambda})(1 - \tau^2/4)^{1/2} \in D_2^*$ (see Fig. 4).

6.3 Multi-well energy: $m = 2, \quad \Lambda u = \frac{1}{2}u^4_x + \frac{1}{2}cu^2_x$

Let $U_a = \{ u \in L^2(0, 1) \mid \Lambda u \in L^2[0, 1], \ u(0) = 0 \}$. The primal problem $(P_m)$ in this case is a non-convex variational problem with two parameters,

$$P_3(u) = \int_0^1 \frac{1}{2}E\left(\frac{1}{2}u^4_x + \frac{1}{2}cu^2_x - \mu\right)^2 \, dx - \int_0^1 fu \, dx \to \min \ \forall u \in U_a.$$  \hspace{1cm} (129)

For the given distributed parameter $c(x)$ and $\mu \in \mathbb{R}$, the energy density

$$V_\mu(p(\epsilon)) = \frac{1}{2}E\left(\frac{1}{2}\epsilon^4 + \frac{1}{2}\epsilon^2 - \mu\right)^2$$

has at most four wells (see Fig. 5). Then non-convex variational problem with a multi-well energy appears frequently in phase transitions and shape memory alloys (see, for example, Bubner 1996; Falk 1990). From the point of view of traditional relaxation methods,