PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

On Dual Approximation Principles and Optimization in Continuum Mechanics

M. J. Sewell

Phil. Trans. R. Soc. Lond. A 1969 **265**, 319-351
doi: 10.1098/rsta.1969.0059

ON DUAL APPROXIMATION PRINCIPLES AND OPTIMIZATION IN CONTINUUM MECHANICS

By M. J. SEWELL
Department of Mathematics, University of Reading
(*Communicated by R. Hill, F.R.S.—Received* 10 *January* 1969)

CONTENTS

A unified expression of some of the boundary value problems of continuum mechanics is developed. A central role is given to the notion of a Legendre dual transformation in displaying the simple analytical structure of each problem considered. A systematic method of deriving reciprocal variational principles is described. General boundary value problems governed by inequalities as well as equations are then considered. Convexity of the dual functions related by the Legendre transformation is shown to be the basis of uniqueness theorems and extremum principles. Attention is drawn to the relevance of the literature on mathematical programming theory. Many examples are given, involving new or recent results in elasticity, plasticity, fluid mechanics and diffusion theory.

1. Introduction

This paper is an exploration of some of the common ground which evidently exists between continuum mechanics and mathematical programming or optimization; and also of the problem of finding a unifying expression of some of the special boundary value problems of continuum mechanics.

Methods of linear and nonlinear programming have been developed rapidly in recent years, and a convenient source of many references to the detailed work in nonlinear programming is the symposium of papers edited by Abadie (1967). Rationalization of the approach to constrained minimization problems in various previously disconnected areas has recently been demonstrated by Canon, Cullum & Polak (1966). Much of the stimulus for such developments in mathematical programming has been found in the fields of control theory and econometrics in the wide sense.

By contrast, apart from certain topics in the theory of structures, there seems to have been little awareness that the existence of such a highly developed body of theory on optimization might be of value in the search for approximate solutions of the boundary value problems of continuum mechanics. Reasons for this situation are not hard to find. The 'constraints' of continuum mechanics, by which we mean here *all* of the governing equations and boundary conditions, usually have the form of equalities rather than inequalities, whereas it is inequalities which characteristically appear in programming theory. Also, the optimization technique in mechanics is often facilitated by the fact that the required extremum (i.e. minimum or maximum) is known to be a *stationary* extremum in a certain class, so that classical variational methods are available. Moreover, there are problems of continuum mechanics whose solutions have a variational characterization, but no extremum property; and of course sometimes no variational principle exists either.

We remark that the first two of the foregoing three reasons should not of themselves be regarded as disqualifying an approach via programming theory, since it is common practice there to replace any equation $M = 0$ (say) which happens to occur by the two inequalities $M \geqslant 0$ and $-M \geqslant 0$. And in any case boundary value problems of continuum mechanics do sometimes involve inequalities, as we illustrate below in various contexts. It is for future investigations to show whether the procedures which have been developed in the specialist optimization literature can be applied systematically and with profit to those problems of continuum mechanics which admit extremum principles (linear programming in plastic limit analysis is one existing example, cited in § 3 (*c*) here); and in particular to show whether any improvement over existing variational techniques is afforded in the special cases when the extrema are stationary.

The third reason above takes one further from the optimization context towards other types of approximation methods such as orthogonalizing or averaging approaches which we shall not discuss here (cf. Hill (1963 *a*) and Whitham (1967, § 4) for discussions in the contexts of solids and fluids respectively).

The genesis of the present investigation was in a comparison of papers by Moreau (1967) and Hill (1956). Moreau examines the boundary value problem in perfect fluid theory described at the end of § 6, and he describes the extremum principles there as infinite-dimensional quadratic programming problems. Also, in the holonomic case of the problem of § 4, Moreau (1966 *a*) remarks that the Kuhn–Tucker 1951 theorem of programming theory can be used to get one of the extremum principles. By contrast, Hill (1956) proposes a systematic procedure for the construction of uniqueness theorems and extremum principles in different branches of continuum mechanics. Hill's proposal gives a central role to the notions of convexity and duality, and was apparently independent of work in the programming field—which is now well known to be intimately concerned with these concepts.

In § 2 here these threads are drawn together in a development from first principles, slanted towards mechanics (continuum or discrete), which enables us to avoid an excursion into the specialized jargon of programming theory. Such an approach seems necessary since few mathematicians are widely experienced in both fields. The roles of duality, adjointness in the governing equations, variational principles, convexity, unilateral and bilateral conditions, uniqueness theorems, extremum principles and passive variables are explored in sequence. A simple matrix notation is used, and one feature is a schematic representation of the permitted inequalities. Several distinct types of boundary value problem are brought within the same analytical framework.

The range of the general theory of § 2 is exemplified in § 3 (see table 1 particularly). Some known results for elastic and plastic solids and perfect fluids, most of them very recent, are summarized. New results suggested by the analytical structure of § 2 are proved in a variety of contexts in §§ 3 to 6, and also in § 2 itself. In particular the systematic approach there to reciprocal variational principles via a free variational principle seems far from well known, and the discussion of § 2 (vii) (related to the diffusion equation, and including the construction of a convex function from a saddle surface) is original.

Parts of the theory presented here can be approached via the notion of function-space and the 'method of the hypercircle' (when variational principles can be strengthened into extremum principles) or the 'pseudo-hypercircle' (when they cannot). But we have not followed this route, preferring the more direct task of seeking a unifying analytical structure (§ 2). The contrast between the two approaches is reinforced by the noteworthy fact that there sems to be no mention in Synge's book (1957) of the notion of convexity.

2. Analytical structure

In this section we explore the basic analytical structure which underlies the type of results which we shall exemplify in this paper in a wide variety of problems. In doing this we deliberately use a non-commital notation which is not too closely identified with any one particular field of application. This facilitates an overall view. We aim to give a very simple and easily intelligible account of some of the main theoretical features, although we do not attempt to cover all variants of detail.

(i) *Legendre dual transformation*

A given continuous scalar function

$$Y = Y(y_i, u_\alpha) \tag{2.1}$$

of n variables y_i $(i = 1, \ldots, n)$ and m variables u_α $(\alpha = 1, \ldots, m)$, having continuous first derivatives, is used to relate to them another set of n variables x_i $(i = 1, \ldots, n)$ by the transformation

$$x_i = \partial Y / \partial y_i. \tag{2.2}$$

This transformation is assumed single-valued at each given set of values of the u_α, and it is reversible, at least locally, if its Jacobian exists and is non-zero, i.e. if

$$\left| \frac{\partial^2 Y}{\partial y_i \partial y_j} \right| \neq 0. \tag{2.3}$$

The inverse may then be written as $\qquad y_i = \partial X / \partial x_i \qquad (2.4)$

in terms of a scalar function $\qquad X = X(x_i, u_\alpha), \qquad (2.5)$

which is defined by inserting the inverse of (2.2) into the right-hand side of

$$X = x_i y_i - Y. \tag{2.6}$$

(It will be understood here and later that summation is to be carried out over the full range of values of any repeated Latin or Greek suffix, and that an isolated suffix can take each of its values in turn.) This Legendre dual transformation is of course completely symmetrical, in that we could have started with (2.5) as given and arrived at (2.2) when (2.3) holds via a definition (2.6) of Y. The u_α play the role of passive variables in the transformation, and have the property

$$\frac{\partial X}{\partial u_\alpha} = -\frac{\partial Y}{\partial u_\alpha} \tag{2.7}$$

in terms of derivatives of (2.5) on the left and of (2.1) on the right.

(ii) *Governing equations*

Now suppose that we are dealing with a problem whose equations impose certain *further* conditions on the y_i, x_i and u_α. We consider first the case in which these further conditions fall into two groups of m and n conditions expressible, for example, as

$$L_{\alpha i} x_i = -\partial Y/\partial u_\alpha \tag{2.8}$$

and

$$y_i = \tilde{L}_{i\alpha} u_\alpha \tag{2.9}$$

respectively. Here the $L_{\alpha i}$ are the elements of an assigned matrix or linear differential operator L, and the $\tilde{L}_{i\alpha}$ are the elements of its adjoint operator $\tilde{L}$. The derivatives in L and $\tilde{L}$ are with respect to a field of k independent variables $\tau_1, \ldots, \tau_k$ spanning the space in which the problem is defined (for example, $k = 4$ includes the case of one time variable with three Euclidean spatial or material coordinates). The x_i, y_i and u_α are the dependent field variables in such a case. Actually the x_i and y_i may arise not only as row or column matrices, but also as rectangular or square matrices (e.g. as stress or strain) having a total of n elements. Even the u_α and $\partial Y/\partial u_\alpha$ might arise as rectangular matrices having a total of m elements. We include such rearrangements by supposing that L and $\tilde{L}$ can also be rearranged to allow (2.8) and (2.9) to be rewritten in the symbolic form

$$Lx = s \quad \text{with} \quad s_\alpha \equiv -\partial Y/\partial u_\alpha, \tag{2.10}$$

$$y = \tilde{L}u, \tag{2.11}$$

where juxtaposition implies operation according to the rules of matrix multiplication. Legendre transformations having passive variables different from those which appear on the right of (2.9) and (2.11) are considered later.

The adjoint property is expressed here by the requirement that, for every u and x,

$$\int (uLx - x\tilde{L}u)\,\mathrm{d}\tau = -\int uNx\,\mathrm{d}\sigma \tag{2.12}$$

for any k-dimensional region τ with bounding surface σ, where N is a matrix operator which is assigned on σ (and is independent of u and x). The integrands in (2.12) are scalars evaluated by the rules of matrix multiplication and contraction. The context will show where the transposes of u and x are required, so that separate notation for these transposes is not introduced here. For example, when

$$u_\alpha L_{\alpha i} x_i - x_i \tilde{L}_{i\alpha} u_\alpha = -\operatorname{div} \boldsymbol{F}(u, x) \tag{2.13}$$

for every u and x, where $\boldsymbol{F}$ is a τ-space vector bilinear in u and x and div indicates divergence in τ-space (cf. Lanczos 1961, equation (4.17.2)), then (2.12) is a consequence of Poincaré's generalization of the divergence theorem (see Ericksen 1960, p. 816), and N represents the normal to σ.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

We shall later confine attention to cases in which uN involves no summation. This is illustrated by the cases when x is rectangular, and when $m = 1$, so that

$$uNx = u_\alpha N_i x_{i\alpha} \quad \text{or} \quad uNx = uN_i x_i \tag{2.14}$$

respectively.

The case when L and $\tilde{L}$ themselves are merely matrices is included by formally setting $k = 0$, and the right-hand side of (2.13) disappears. The adjoint is then just the transpose, so that

$$\tilde{L}_{i\alpha} = L_{\alpha i} \tag{2.15}$$

replaces (2.13) and (2.12), and can be inserted into (2.11). Of course, in writing down (2.8) and (2.9) we have destroyed the symmetry between the x_i and y_i which (when (2.3) held) we had up to that point, and this will be reflected in applications requiring particular interpretations for the x_i and y_i. One interpretation illustrated in § 3 is when (2.10) and (2.11) represent equilibrium equations and strain-displacement equations respectively.

(iii) *Free variational principles*

We define next a 'canonical action' functional

$$A[x, y, u, Lx] = \int [xy - uLx - Y(y, u)]\, \mathrm{d}\tau. \tag{2.16}$$

In view of (2.12) the first variation of this for *unrelated* small variations δx, δy and δu is

$$\delta A = \int [\delta x(y - \tilde{L}u) + (x - \partial Y/\partial y)\delta y - \delta u(Lx - s)]\, \mathrm{d}\tau + \int uN\delta x\, \mathrm{d}\sigma. \tag{2.17}$$

The variational principle

$$\delta A - \int uN\delta x\, \mathrm{d}\sigma = 0 \tag{2.18}$$

will therefore yield (2.2), (2.10) and (2.11) as natural conditions without making *any* prior assumption about the relation between x, y and u. We could therefore use (2.18) as a starting-point from which to deduce (2.2) (instead of assuming it), and then add restrictions such as (2.3) or the convexity (mentioned below) subsequently. Various specializations and generalizations of (2.18) are possible, as the sequel will show. An alternative to (2.16) would be to begin with the integral of $xy - x\tilde{L}u - Y(y, u)$ over τ. In *coupled* problems, such as the air/membrane vibration problem referred to in § 3(*h*), A can be a sum of integrals over regions of differing dimensions.

The assertion (2.18) is a free variational principle formulated so that *all* of its natural conditions apply throughout τ, and therefore none of them apply only on the boundary σ. There now arises the question of what boundary conditions on the x and u does the form of uNx allow boundary integrals to be inserted on the right-hand side of (2.18) so that the result takes the more viable form

$$\delta(A - B) = 0. \tag{2.19}$$

Here in general one would expect B to be an integral over the $(k-1)$-dimensional boundary σ, and of most interest are the forms (such as (2.14)) of uNx which allows the boundary conditions on σ to emerge as natural conditions on σ of (2.19) (over and above those in τ of (2.18)). The most common and easiest case in which to establish (2.19) from (2.18) is when certain orthogonal linear combinations of u (and/)or x have assigned values at every point of σ. For example, suppose (2.14) or a similar form of uNx holds, and

$$u = h \quad \text{on a part} \quad \sigma_u \quad \text{of} \quad \sigma, \tag{2.20}$$

and

$$Nx = T \quad \text{on a part} \quad \sigma_x \quad \text{of} \quad \sigma \tag{2.21}$$

for assigned quantities h and T, where $\sigma_u + \sigma_x = \sigma$ to begin with.

These boundary conditions satisfy and also emerge from the principle

$$\delta A - \int uN\delta x \mathrm{d}\sigma = -\int (u-h)\, N\delta x \,\mathrm{d}\sigma_u + \int \delta u (Nx - T)\mathrm{d}\sigma_x. \tag{2.22}$$

This is obtained by augmenting (2.18) in an obvious way which also achieves the alternative form

$$\delta\left[A - \int u(Nx-T)\,\mathrm{d}\sigma_x - \int hNx\,\mathrm{d}\sigma_u\right] = 0, \tag{2.23}$$

thus exemplifying (2.19). Such simple boundary conditions therefore take care of themselves (other considerations apart), at least on finite boundaries. Prescribed displacement and/or surface load typify this case in solid mechanics; other examples may be inferred from § 3.

It can happen, however, that a more complicated relationship between u and x is assigned on part of σ. Such a relationship may be nonlinear and it may also involve certain τ-derivatives of u and x. The known criteria for the existence and determination of B in such cases seem far from complete. However, some conclusive information is contained in a general study by Sewell (1967) of nonlinear configuration-dependent loading, following upon some sufficient conditions obtained by Hill (1962) in the linearized problem. Uniform normal pressure loading maintained through a large deflexion of the loaded surface is a principal nonlinear example in which B can exist in this more complicated case (Sewell 1967, equation (89); 1965, equation (32)).

Suppose then that (2.19) is of such a form (e.g. (2.23) with (2.14)) that *all* the governing equations in τ and the boundary conditions on σ are recoverable from it as natural conditions after unconstrained and independent small variations of x, y and u are applied. Then reciprocal or dual variational principles are obtainable from (2.19) (as well as from (2.18)) in various ways. The general procedure requires that some (or none) of the natural conditions are assumed to hold from the outset, and then that complementary sets of what remains of these original natural conditions be imposed in turn on (2.19) before the variations are carried out. A pair of constrained variational principles is obtained whose natural conditions complement each other in the sense that, when taken together with the overall constraints applied at the outset, they make up the entire set of the original natural conditions of (2.19) itself. A specific new illustration of this procedure in compressible fluid mechanics is given by Sewell (1963 *a*). Of most interest are likely to be those reciprocal principles which verify that complementary extremum principles are analytic (when extremum principles exist at all), or at least those which involve variations among only *one* set of variables in each individual principle. To get extremum principles requires more specific assumptions about Y than we have made hitherto.

(iv) *Convex functions*

Hill has proposed (1956) a general automatic procedure for the construction of uniqueness theorems and associated extremum principles in continuum mechanics, and it is an extension of this procedure which we require next. Broadly speaking this procedure requires that once the governing equations have been cast into the form (2.2) with (2.10) and (2.11), the function Y should be a convex function of the y_i in some suitable sense.

We require in §§ (v) and (vi) below the fairly common case in which Y depends only linearly on the u_α, with coefficients $-s_\alpha$ which are assigned and therefore independent of the y_i, so that

$$Y(y_i, u_\alpha) = U(y_i) - s_\alpha u_\alpha, \tag{2.24}$$

where $U(y_i)$ is an assigned function of the y_i alone. The interpretations of (2.10) in § 3 show that *such* s can often represent assigned body force. Then the $\partial Y/\partial y_i$ are independent of the u_α.

We shall use the prefix Δ applied to a function to indicate the ordered difference of the function-values corresponding to any two distinct values of its argument. Suppose that the general $Y(y_i, u_\alpha)$ has the property that, for every fixed u and any $\Delta y \neq 0$,

$$\int [\Delta Y - (\partial Y/\partial y)\,\Delta y]\,\mathrm{d}\tau > 0, \tag{2.25}$$

where the matrix $\partial Y/\partial y$ is evaluated at the y-point corresponding to the 'minus member' of the difference $\Delta y (=$ 'plus member' less 'minus member').

There is no loss of generality in reversing the roles of the two members of Δy in (2.25), and if we do this and add the result to (2.25) we get

$$\int \Delta(\partial Y/\partial y)\,\Delta y\,\mathrm{d}\tau > 0. \tag{2.26}$$

The second mean value theorem applied to the present Y shows that for every fixed u

$$\Delta Y - \frac{\partial Y}{\partial y_i}\Delta y_i = \tfrac{1}{2}\Delta y_i \Delta y_j \overline{\frac{\partial^2 Y}{\partial y_i \partial y_j}}, \tag{2.27}$$

where the bar indicates evaluation of these second derivatives at some y-point between the two members of Δy. If these two members are separated by a surface of discontinuity of the second derivatives, the bar might have to imply a value within the range of the 'jump' at the discontinuity. It follows that, if

$$\frac{\partial^2 Y}{\partial y_i \partial y_j} \text{ is positive definite at every point,} \tag{2.28}$$

then $Y(y_i, u_\alpha)$ is a convex function of the y_i at each u_α. For (2.27) with (2.28) implies that the left-hand side of (2.27) is positive for every $\Delta y \neq 0$, and the geometrical interpretation of this is that the surface (2.1) lies entirely on one side of its tangent plane at any y-point (for each fixed u), which illustrates the convexity. Condition (2.28) (which incidentally implies (2.3)) is therefore sufficient for convexity (defined here by requiring (2.27) to be positive for every $\Delta y \neq 0$), but is somewhat stronger than necessary—for example, isolated zeros of $\partial^2 Y/\partial y_i \partial y_j$ could be admitted. In any case it is the definition of 'overall convexity' (2.25) which will be adequate for our purpose, and this could permit (2.27) to be $\geqslant 0$ and even < 0 in part of τ provided it is > 0 in 'most' of τ. A discontinuity of $\partial^2 Y/\partial y_i \partial y_j$ along the join through the origin of two different quadratic functions Y with continuous slope is a fundamental property of, for example, plasticity theory, where (2.2) represents the constitutive equations for the boundary value problem of quasi-static infinitesimal displacement (as illustrated in §3 (*b*)). Hill gives a proof (1961 *a*, §4; 1962, §5) indicating that (2.26) implies (2.25) for continuous Y with continuous first derivatives and piecewise continuous second derivatives, so that (2.25) and (2.26) are completely equivalent for such functions. All the properties of Y stated in this paragraph 'for each fixed u' also apply to $U(y_i)$ without such qualification about u being required—such properties of U also ensue from a widened definition of convexity in which the restriction $\Delta u_\alpha \equiv 0$ is dropped from (2.25).

When (2.2) and (2.3) hold, so that the full Legendre dual transformation is available, use of (2.2), (2.4) and (2.6) in the integrand of (2.5) shows that (2.25) is then equivalent to

$$\int [\Delta X - (\partial X/\partial x)\,\Delta x]\,\mathrm{d}\tau > 0 \tag{2.29}$$

for every fixed u and any difference $\Delta x \neq 0$. Here again the matrix $\partial X/\partial x$ is evaluated at the x-point corresponding to the 'minus member' of the difference Δx. Thus the overall convexity of either function generating the Legendre transformation implies the overall convexity of its dual function. An alternative to (2.29) is

$$\int \Delta(\partial X/\partial x)\,\Delta x\,\mathrm{d}\tau > 0, \tag{2.30}$$

by the same argument which relates (2.26) to (2.25).

In the special case (2.24) we are here dealing with a function

$$X(x_i, u_\alpha) = U_c(x_i) + s_\alpha u_\alpha, \tag{2.31}$$

which is linear in the u_α with assigned coefficients $s_\alpha = \partial X/\partial u_\alpha$ independent of the x_i, and where $U_c(x_i)$ is an assigned function of the x_i alone. Equations (2.29) and (2.30) also hold with X replaced by U_c and the u-qualification dropped. Equation (2.31) follows from (2.24) with $xy - U = U_c$.

It may of course happen that we are given a function (2.5) satisfying (2.29) but for which (2.4) cannot be inverted, so that Y in the form (2.1) does not exist. In that case we can rewrite (2.29) in the form

$$\int [\Delta(xy - X) - x\Delta y]\, d\tau > 0, \tag{2.32}$$

where we have used (2.4). Clearly (2.32) would become (2.25) if the transformation (2.4) were reversible.

(v) *Uniqueness theorems under bilateral and unilateral conditions*

Now consider the difference Δy of any two y-distributions which are representable in the form (2.11), so that such $\Delta y = \tilde{L}\Delta u$. These have the property that

$$\int \Delta x \Delta y\, d\tau = \int \Delta u L \Delta x\, d\tau + \int \Delta u N \Delta x\, d\sigma \tag{2.33}$$

by (2.12). Assuming (2.2) and (2.26), a *reductio ad absurdam* argument shows that there cannot be more than one solution for y of any boundary value problem whose definition is completed by conditions which would make the right-hand side of (2.33) non-positive. As an example, one such problem is completed by adding (2.10) (with assigned s) in τ and the boundary conditions (2.20) and (2.21) on σ. For this problem each integral on the right of (2.33) would have to be actually zero. This problem is completed by 'bilateral conditions' (equalities), but uniqueness of y also follows when the problem is completed by 'unilateral conditions' (inequalities) of a certain type on either or both of τ and σ, or by a suitable combination of the two types of conditions on different parts of τ and of σ. The transition from bilateral to appropriate unilateral conditions can be graphically illustrated schematically (recall the matrix multiplication and contraction to inner products in (2.33)) as in figures 1 and 2. Figure 1 (a) represents the matrix equation (2.10) with assigned s. Figure 1 (b) represents unilateral conditions of the type

$$Lx \geqslant s, \tag{2.34}$$

and

$$u \geqslant 0, \tag{2.35}$$

with

$$u(Lx - s) = 0. \tag{2.36}$$

(A matrix inequality refers to every element of the matrix.) Such inequalities are illustrated by the unilateral internal kinematical constraints analysed in §§4 to 6. Figure 2 represents a replacement of (2.20) (figure 2(b)) and (2.21) (figure 2 (a)) on at least a part σ_c of the original $\sigma_x + \sigma_u$ by unilateral conditions which have the properties

$$Nx - T \geqslant 0 \tag{2.37}$$

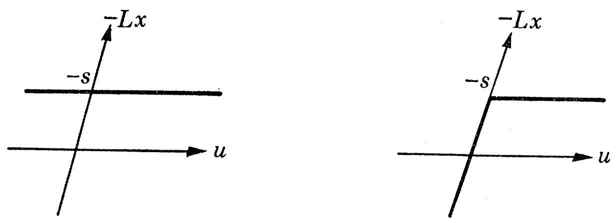
and

$$u - h \geqslant 0, \tag{2.38}$$

with

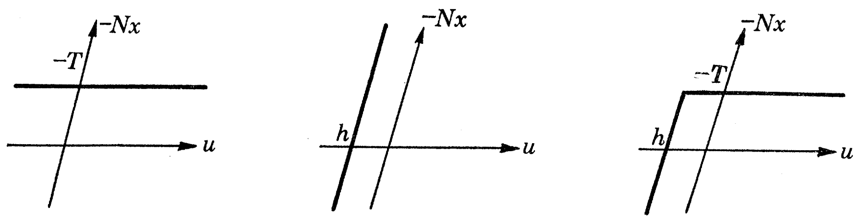
$$(u-h)(Nx - T) = 0. \tag{2.39}$$

Such unilateral surface constraints are illustrated by the passive constraints described in §3. It is easy to show that (2.34) to (2.36) make the τ-integrand in (2.33) non-positive, and that (2.37) to (2.39) make the σ-integrand in (2.33) non-positive on σ_c. Therefore, for example, a boundary value problem requiring the solution of (2.11) and (2.34) to (2.36) subject to (2.37) to (2.39) cannot involve more than one y-distribution when (2.2) and (2.26) hold. The x-distribution is then unique by the single-valuedness of (2.2) (which does not contain passive variables in the present problem). But the u-distribution will not be unique if the boundary conditions do not exclude trivial non-zero solutions of $\tilde{L}u = 0$ (however they often do, as in the Dirichlet problem of potential theory, and in the exclusion of rigid-body solutions to the equations of classical elasticity by suitable boundary constraints on the displacement). The same conclusions follow if figures 1 (*b*) and/or 2 (*c*) enclose the opposite quadrant, for which both inequalities (2.34) and (2.35), and/or both (2.37) and (2.38), are reversed.



(*a*) (2.10) with assigned s (*b*) unilateral form of (*a*)

FIGURE 1.



(*a*) $Nx = T$ on σ_x (*b*) $u = h$ on σ_u (*c*) unilateral on σ_c

FIGURE 2.

An alternative approach to the uniqueness question can exploit (2.30) instead of (2.26). Consider the difference Δx of any two x-distributions which satisfy either (2.10) or (2.34) to (2.36). Then the associated Δy obtained via (2.4), and any Δu (satisfying (2.35) and (2.36) if these are applicable) have the property that

$$\int \Delta x \Delta y \, d\tau \leqslant \int \Delta x \Delta (y - \tilde{L}u) \, d\tau + \int \Delta u N \Delta x \, d\sigma \tag{2.40}$$

by (2.12) and (2.34) to (2.36). Then assuming (2.30), a *reductio ad absurdam* argument shows that there cannot be more than one solution x of any boundary value problem whose definition is completed by conditions which make the right-hand side of (2.40) non-positive. We do not pursue here the modification of (2.11) to unilateral conditions which this alternative suggests. The comparative strengths of criteria of uniqueness based on (2.26) and (2.30) depend on the particular problem concerned. A comparison for the classical elastic solid subject to bilateral equilibrium, strain-displacement and boundary conditions has been made by Hill (1961 *b*). In practice the definitions (2.25) and (2.29) of overall convexity may often be weakened without serious loss. When all the conditions are bilateral, for example, we may require (2.25) to hold not for all $\Delta y \neq 0$, but for the smaller class of all $\Delta y \neq 0$ computed from the solutions of (2.11) and (2.20);

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

and we may require (2.29) to hold not for all $\Delta x \neq 0$, but for the smaller class of all $\Delta x \neq 0$ computed from the solutions of (2.10) and (2.21).

(vi) *Extremum principles*

The unique solutions which are so guaranteed by (2.26) or (2.30) can be characterized by extremum principles which provide upper and lower bounds to the value of a certain scalar quantity in the actual solution. These extremum principles follow from appropriate interpretations of the alternative statements of overall convexity contained in (2.25) and (2.29) (or (2.32)). In the following illustrations we regard the 'minus members' associated with the differences in (2.25) and (2.29) or (2.32) as having *all* the properties of an actual solution, and they will be undesignated (in contrast to the 'plus members' which are designated below by either * or †). To illustrate the derivation of extremum principles the properties of an actual solution will now be taken as follows: either (2.10) or (2.34) to (2.36) in τ (with assigned s), together with (2.11) in τ; either (2.2) or (2.4) or both in τ; and $\sigma = \sigma_x + \sigma_u + \sigma_c$ where (2.20) holds on the new σ_u, (2.21) on the new σ_x and (2.37) to (2.39) on σ_c.

With these data and $\Delta x = x^* - x$ inserted into (2.29) we find with the aid of (2.12) that, since $\Delta X = \Delta(U_c + su) = \Delta U_c$ by hypothesis, (2.29) may be written

$$\int \Delta U_c \, d\tau - \int hN\Delta x \, d(\sigma_u + \sigma_c) > \int u(Lx^* - s) \, d\tau + \int (u-h)(Nx^* - T) \, d\sigma_c. \tag{2.41}$$

Here x^* is an approximate solution required only to satisfy (2.10) or (2.34) in τ, (2.21) on σ_x and (2.37) on σ_c. Hence the righthand side of (2.41) is zero when the conditions in τ and on σ are all bilateral, and non-negative if they are partially or wholly unilateral. We therefore have the extremum principle

$$\int U_c(x^*) \, d\tau - \int hNx^* \, d(\sigma_u + \sigma_c) > \int U_c(x) \, d\tau - \int hNx \, d(\sigma_u + \sigma_c). \tag{2.42}$$

It is worth noting that the statement (2.42) of this principle is the same regardless of whether it is (2.10) or (2.34) to (2.36) which holds in τ. (There is no connexion between the two uses of a subscript c in U_c and σ_c).

With the same data for an actual solution and $\Delta y = y^\dagger - y$ inserted into (2.32), we find with the aid of (2.12) and the hypothesis $\Delta X = \Delta U_c$ that (2.32) may be written

$$\int \Delta(xy - U_c - su) \, d\tau - \int T\Delta u \, d(\sigma_x + \sigma_c) > \int u^\dagger(Lx - s) \, d\tau + \int (u^\dagger - h)(Nx - T) \, d\sigma_c. \tag{2.43}$$

Here $y^\dagger = \tilde{L}u^\dagger$ is an approximate solution required only to satisfy (2.11) (and (2.35) at points of τ where the unilateral conditions apply) in τ, (2.20) on σ_u and (2.38) on σ_c. Hence the right-hand side of (2.34) is zero when the conditions in τ and on σ are all bilateral, and non-negative if they are partially or wholly unilateral. We therefore have another extremum principle by dropping the right-hand side of (2.43), and when in addition (2.4) is reversible so that $U = xy - U_c$ this may be written (and then alternatively obtained from (2.25)) as

$$\int [su - U(y)] \, d\tau + \int Tu \, d(\sigma_x + \sigma_c) > \int [su^\dagger - U(y^\dagger)] \, d\tau + \int Tu^\dagger \, d(\sigma_x + \sigma_c). \tag{2.44}$$

It is worth noting that the statement (2.44) of this principle is also the same regardless of whether it is (2.10) or (2.34) to (2.36) which holds in τ.

If we denote the minimum which the actual solution provides in (2.42) by $\min_{x^*}$, and the maximum which it provides in (2.44) by $\max_{u\dagger}$, it is easily shown that

$$\min_{x^*} = \max_{u\dagger} - \int hT \, d\sigma_c. \tag{2.45}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Therefore the extremum principles are complementary, in the sense that they provide upper and lower bounds to the same quantity in the actual solution, whenever $\min_{x^*} \geqslant \max_{u^\dagger}$, that is whenever

$$\int hT\,\mathrm{d}\sigma_c \leqslant 0. \tag{2.46}$$

The solution values themselves of the bounded quantities coincide when equality holds in (2.46). This happens whenever the matrices h and T are orthogonal to one another, in either a pointwise or an overall sense, on the part σ_c of the boundary where unilateral conditions are specified; and in particular where $h = 0$ or $T = 0$ on σ_c, or when $\sigma_c = 0$ (the last case permits only the bilateral conditions on σ).

When all the conditions are bilateral both extrema are analytic. This is indicated by the right-hand side of (2.27), which suggests that what has been given away to produce the inequalities in (2.25) and (2.29) and therefore in (2.42) and (2.44) is of second order in the differences Δy and Δx (supposed small for this purpose). It can be proved from (2.23) by applying the procedure for deriving reciprocal variational principles which has already been described. We have only to impose (2.2) and (2.3), and then either (2.11) and (2.20) to give

$$\delta\{\int [su^\dagger - U(y^\dagger)]\,\mathrm{d}\tau + \int Tu^\dagger\,\mathrm{d}\sigma_x\} = 0,$$

or (2.10) and (2.21) to give $\quad \delta\{\int U_c(x^*)\,\mathrm{d}\tau - \int hNx^*\mathrm{d}\sigma_u\} = 0.$

When the conditions are partly unilateral an extra amount represented by the right-hand sides of (2.41) and (2.43) is given away to get the extremum principles. In the absence of a guarantee that these extra amounts are not of first order in Δx or Δu respectively, we are not entitled to suppose that the extrema are stationary when unilateral conditions are present.

(vii) *Passive variables occurring nonlinearly*

We now drop the restriction (2.31) that $X(x_i, u_\alpha)$ depends only linearly on the u_α, and admit a general dependence on all the variables in (2.5). We wish to discuss an example in which we suppose from the outset that the governing equations of the problem in τ are

$$Lx = \partial X/\partial u, \tag{2.47}$$

$$\tilde{L}u = \partial X/\partial x \tag{2.48}$$

in matrix form (an explicit example is $L_{\alpha i}x_i = \partial X/\partial u_\alpha$, $\tilde{L}_{i\alpha}u_\alpha = \partial X/\partial x_i$). These equations are deducible from the definition (2.4) with (2.11), and (when the Legendre transformation beginning with (2.5) is reversible) from (2.7) with (2.10). They can still be postulated without assuming that (2.4) are reversible, and they can be regarded as derivable from a form of (2.18) in which instead of (2.16) we use

$$A[x, u, Lx] = \int [X(x, u) - uLx]\,\mathrm{d}\tau. \tag{2.49}$$

Reciprocal variational principles can be obtained by choosing u to be such that (2.47) is satisfied for all x (which can certainly be done if $|\partial^2 X/\partial u_\alpha\,\partial u_\beta| \neq 0$), whence $\delta\int(X - u\,\partial X/\partial u)\,\mathrm{d}\tau - \int uN\delta x\,\mathrm{d}\sigma = 0$ implies (2.48) from the coefficient of δx alone; and, on the other hand, by choosing x to be such that (2.48) is satisfied for all u (which can certainly be done if $|\partial^2 X/\partial x_i\,\partial x_j| \neq 0$, which also ensures that the original Legendre transformation *is* reversible), whence $\delta\int(X - x\,\partial X/\partial x)\,\mathrm{d}\tau + \int \delta uNx\,\mathrm{d}\sigma = 0$ implies (2.47) from the coefficient of δu alone.

To strengthen these principles into complementary extremum principles it is sufficient that (apart from the question of boundary conditions) the function $X(x, u)$ shall have the following properties. For the first extremum principle the function $X - u\,\partial X/\partial u$ of x and Lx which is obtained via the inverse of (2.47) must be convex (or concave) with respect to *these* variables. It may be shown that for *this* function the expression upon which the definition of convexity is based (a generalization of the integrand of (2.29)) may be written

$$\Delta X - \frac{\partial X}{\partial x}\bigg|_{-} \Delta x - \frac{\partial X}{\partial u}\bigg|_{+} \Delta u \tag{2.50}$$

in terms of the partial derivatives at x_- and x_+ of the function $X(x, u)$ as originally given, where the prefix Δ denotes the difference of quantities associated (via (2.47)) with the points x_- and x_+ in the order typified by $\Delta x = x_+ - x_-$. So that the above-mentioned function of x and Lx is convex (concave) when (2.50) is positive (negative) for every distribution of Δx and $L\Delta x$ which are not all zero, in which event the first variational principle may be strengthened to an extremum principle for suitable boundary conditions. Likewise the second extremum principle requires that the function $X - x\,\partial/X\partial x$ of u and $\tilde{L}u$ obtained via the inverse of (2.48) must be concave (or convex) with respect to *these* variables. It may be shown that for *this* function the expression upon which the definition of convexity is based (a generalization of the left-hand side of (2.27)) may be written

$$\Delta X - \frac{\partial X}{\partial x}\bigg|_{+} \Delta x - \frac{\partial X}{\partial u}\bigg|_{-} \Delta u \tag{2.51}$$

in terms of the partial derivatives at u_- and u_+ of the function $X(x, u)$ as originally given, where the prefix Δ denotes the difference of quantities associated (via (2.48)) with the points u_- and u_+ in the order typified by $\Delta u = u_+ - u_-$. So that the above function of u and $\tilde{L}u$ is concave (convex) when (2.51) is negative (positive) for every distribution of Δu and $\tilde{L}\Delta u$ which are not all zero. From this the second variational principle may for suitable boundary conditions be strengthened to an extremum principle complementary to the first one in that it provides a lower (upper) bound to the quantity bounded above (below) by the first extremum principle.

When the original function $X(x_i, u_\alpha)$ is quadratic with diagonalized second degree terms it is obvious that the above two requirements of convexity and concavity imply that the surface $X = X(x_i, u_\alpha)$ must itself represent a saddle surface convex (or concave) with respect to the x_i and concave (or convex) with respect to the u_α. In such a special case, for which X is given by (3.25), Arthurs (1967) derives upper and lower bounds for the absorption probability associated with neutron diffusion in solids, for the case of a uniform isotropic source within τ.

When either of (2.50) and (2.51) is always positive (or always negative) we can exchange the role of 'plus' and 'minus' members and add, giving (e.g. from (2.50) when negative, or from (2.51) when positive)

$$\Delta u \Delta\left(\frac{\partial X}{\partial u}\right) - \Delta x \Delta\left(\frac{\partial X}{\partial x}\right) > 0. \tag{2.52}$$

If we now suppose that there are at least two distinct possible solutions of (2.47) and (2.48) with differences Δx and Δu, we see from (2.12) that the boundary conditions would have to satisfy

$$\int \Delta u N \Delta x \,\mathrm{d}\sigma < 0. \tag{2.53}$$

Uniqueness therefore follows by *reductio ad absurdam* if the boundary conditions actually make the surface integral in (2.53) positive or zero (it is zero in the neutron diffusion problem. Arthurs does not mention the uniqueness question).

Since this paper was completed, Pomraning has published (1968) a generalization of Courant and Hilbert's notion of involutary and canonical transformations of variational problems.

Many of his equations may be identified with those here (e.g. his (2) and (20) are the dual extremum principles given after (2.49)). The more straightforward development of the present §2 still seems preferable for the class of problems envisaged in §3, although Pomraning's example of a particle transport problem in reactor physics illustrates the role of integral operators. He quotes an example in which 'exact' calculations of a quantity having known upper and lower bounds give a result which (owing to errors in numerical integration) fall *outside* the bounds.

(viii) *Additional passive variables*

Finally we consider the more general Legendre transformations in which passive variables other than the u_α appearing in (2.11) are present. Suppose, for example, that μ is a matrix consisting of extra variables over and above those already considered, and that μ *together with its* τ*-space gradient* $G\mu$ are present as additional passive variables in (2.1) and/or (2.5). Let $\tilde{G}$ denote the adjoint of G ($-\tilde{G}$ will be the div of τ-space). Then with

$$Y = Y(y, u, \mu, G\mu), \tag{2.54}$$

a modified form of (2.18) will yield the previous natural conditions in τ, together with

$$\frac{\partial Y}{\partial \mu} + \tilde{G}\left(\frac{\partial Y}{\partial (G\mu)}\right) = 0 \tag{2.55}$$

as an extra condition in τ from the coefficient of $\delta\mu$, and no natural conditions on σ. Alternatively, with

$$X = X(x, u, \mu, G\mu) \tag{2.56}$$

in (2.49), another suitably modified form of (2.18) will yield

$$\frac{\partial X}{\partial \mu} + \tilde{G}\left(\frac{\partial X}{\partial (G\mu)}\right) = 0 \tag{2.57}$$

as an extra condition in τ. Both (2.55) and (2.57) have the characteristic Euler–Lagrange structure familiar in the calculus of variations. Such variational principles can subsequently be modified to take account of boundary conditions also, in the way previously described. We consider two illustrations of additional passive variables.

In the first illustration we suppose that a function

$$X(x_i, u_\alpha, \mu) = \mu f(x_i) + s_\alpha u_\alpha \tag{2.58}$$

is given, where the s_α are assigned numbers (cf. (2.31)), the function $f(x_i)$ is a given scalar-valued function, and μ is a single additional scalar variable whose τ-space gradient is absent from (2.58). The principle (2.23) with the A of (2.49) and the X of (2.58) has natural conditions

$$\delta u\colon\quad Lx = s \quad \text{in} \quad \tau, \qquad Nx = T \quad \text{on} \quad \sigma_x, \tag{2.59}$$

$$\delta\mu\colon\quad f(x) = 0 \quad \text{in} \quad \tau, \tag{2.60}$$

$$\delta x\colon\ \mu\,\partial f/\partial x = \tilde{L}u \quad \text{in} \quad \tau, \qquad u = h \quad \text{on} \quad \sigma_u. \tag{2.61}$$

Here (2.60) is the present form of (2.57). Reciprocal variational principles can be obtained from this free principle in various ways.

Now suppose that $f(x)$ is a convex function such that

$$\Delta f - (\partial f/\partial x)\,\Delta x > 0 \tag{2.62}$$

for every non-zero $\Delta x = x_+ - x_-$, where $\partial f/\partial x$ is evaluated at x_-. Suppose also that μ is now restricted by

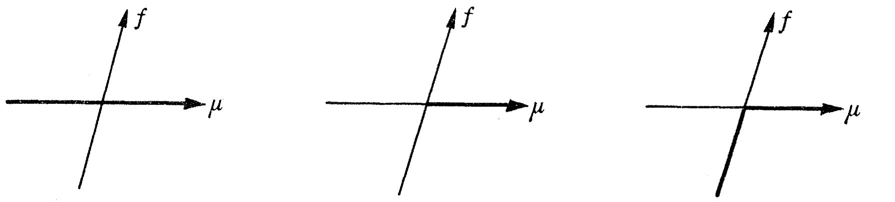
$$\mu > 0. \tag{2.63}$$

There follow the equivalent forms (cf. (2.29) and (2.32))

$$\mu\Delta f > \mu(\partial f/\partial x)\,\Delta x, \tag{2.64}$$

$$-\mu^{\dagger}\Delta f > -\mu^{\dagger}(\partial f/\partial x)|_{+}\Delta x \tag{2.65}$$

for any $\mu > 0$ and $\mu^{\dagger} > 0$.



(*a*) equation (2.60) (*b*) equations (2.60) and (2.63) (*c*) equations (2.69) to (2.71)

FIGURE 3.

Extremum principles strengthening certain of the reciprocal variational principles just mentioned can now be proved by integration of (2.64) and (2.65) over τ, with suitable interpretations of the integrand. If in (2.64) x_{-} is a solution x of all the conditions, and $x_{+} = x^{*}$ (say) is a solution of (2.59) and (2.60) only, then

$$0 = \int \mu[f(x^{*}) - f(x)]\,\mathrm{d}\tau > \int hN\Delta x\,\mathrm{d}\sigma_{u}, \tag{2.66}$$

if μ is the actual solution value. On the other hand, if in (2.65) x_{-} is a solution x of all the conditions, and $x_{+} = x^{\dagger}$ (say) with $\mu^{\dagger}$ and $u^{\dagger}$ is a solution of (2.60) and (2.61) only, then

$$0 = \int \mu^{\dagger}[f(x) - f(x^{\dagger})]\,\mathrm{d}\tau > \int u^{\dagger}s\,\mathrm{d}\tau + \int hNx\,\mathrm{d}\sigma_{u} + \int u^{\dagger}T\,\mathrm{d}\sigma_{x} - \int x^{\dagger}\tilde{L}u^{\dagger}\,\mathrm{d}\tau. \tag{2.67}$$

These principles are complementary and may be stated as

$$\int hNx^{*}\,\mathrm{d}\sigma_{u} < \int hNx\,\mathrm{d}\sigma_{u} < \int x^{\dagger}\tilde{L}u^{\dagger}\mathrm{d}\tau - \int u^{\dagger}T\mathrm{d}\sigma_{x} - \int u^{\dagger}s\,\mathrm{d}\tau. \tag{2.68}$$

Except for the fact that these extrema are analytic, these principles are interpretable as the principle of maximum plastic work (2.66) and its dual (see Hill 1951, and § 3 (*c*) below for more detailed verification). But this analytic property is removed by changing the relations (2.60) and (2.63) between f and μ according to the transition from figure 3 (*a*) through 3(*b*) to 3(*c*). Figure 3 (*c*) may be expressed as

$$f \leqslant 0, \tag{2.69}$$

$$\mu \geqslant 0, \tag{2.70}$$

$$f\mu = 0. \tag{2.71}$$

The left-hand sides of (2.66) and (2.67) are no longer zero but non-positive, and since the extra amounts given away are $\int \mu f(x^{*})\,\mathrm{d}\tau \leqslant 0$ and $\int \mu^{\dagger} f(x)\,\mathrm{d}\tau \leqslant 0$, which are in general bigger than the second order, the extrema cease to be analytic. This argument may be contrasted with the similar arguments based on figures 1 and 2 (which themselves suggested this approach to the maximum work principle because of their superficial similarity with the plastic 'stress–strain' graph). The extra amounts given away here are from the *first* rather than (as previously) the second group of terms in the convexity inequality. A uniqueness theorem for x can also be inferred from the integrated form of the inequality

$$\Delta(\mu\,\partial f/\partial x)\,\Delta x \geqslant 0 \tag{2.72}$$

implied by (2.62) and (2.69) to (2.71).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The second illustration of the role of passive variables contributing equations like (2.55) and (2.57) is in a variational principle which results in the general equations of continuity, motion and energy for the unsteady flow of a compressible inviscid fluid. The discussion of this is left until § 3 (*e*).

Other roles of passive variables in extremum principles are mentioned by Hill (1956), and another approach to extremum principles is described by Ziegler (1963).

3. Illustrations

The purpose of this section is to present some illustrations of the variety of theories having the analytical structure outlined in § 2. Additional illustrations, from a class of initial motion problems, are described in §§ 4 to 6. Attention is confined to some of the most familiar materials, having trivial memory properties, since the possible existence of a unified treatment such as that afforded by § 2 seems far from well appreciated even for such intensively studied solids and fluids—specific materials are too often investigated in isolation from related developments in other continua. The governing equations are assembled for ease of reference in table 1. There now follows a brief explanation of the notation and context of each example. No attempt will be made to list explicitly all the variational principles, uniqueness theorems and extremum principles obtainable under each heading (the reader will be able to infer the approach to these, and the statements of or restrictions upon them, by comparing table 1 with § 2). However, there are certain results whose novelty of content or treatment justify some amplification of the discussion. For brevity we shall leave the possible presence of singular surfaces out of consideration. The notation for each individual example is of course consistent, but the occasional use of the same symbol in *different* contexts need not confuse the reader.

(*a*) *Equilibrium problem of finite elastic strain*

The notation for stress and displacement gradient in table 1 (*a*) is the same as that explained in § 5 below. That is, the u_j are the finite displacement components from a known reference configuration and the s^{ij} are the current values of nominal stress. These tensor components are defined with respect to the reference position of material coordinates θ^i, which may be Cartesian then if desired (in which case all subscripts following a comma are equivalent to partial derivatives). The strain energy per unit reference volume is $U(d_{ij})$, an assigned (but not unrestricted) function of 'strain' variables d_{ij} which are defined by (2.11) to be the displacement gradients $u_{j,i}$. The function $U_c(s^{ij})$ is the complementary energy. Equations (2.10) are the equilibrium equations under assigned body force g^j per unit reference volume (the ρb^j of equations (5.9)), so that we are dealing with an example of the type (2.24) in which the s_α are the body force components g^j. The integrations in (2.12) *et seq.* are over the region τ with bounding surface σ which the body occupies in its reference configuration, which differs in general from the current configuration. The contracted matrix product uNx is $u_j n_i s^{ij}$ in this case, where $\boldsymbol{n}$ is the unit normal to σ outward from τ.

The choice of appropriate variables for setting up this problem has been a matter of some debate in the literature. Here we have chosen to follow Hill (1956, 1957*a*) in using s^{ij} and $u_{j,i}$, because these correspond to the matrices x and $\tilde{L}u$ respectively which the general framework of § 2 requires. Various variational principles which flow from the resulting (2.19) in the ways indicated have been explored by a number of authors, and some review of the literature is given

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY
PHILOSOPHICAL TRANSACTIONS OF

Table 1. Correspondences between general theory of §2 and examples of §§3 to 5

	(a)	(b)	(c)	(d)	(e)	(f)	(g)
$Y(y_i, u_\alpha, \mu, G\mu)$	$U(d_{ij}) - g^j u_j$	$U(d_{ij}) - \dot{g}^j v_j$	—	$L(v_i, q_i, t)$	$-p(\boldsymbol{y}; \alpha, \beta, h, \eta; \nabla h, \nabla\eta)$	$\frac{1}{2}a_{ij}^{-1}(y_i+z_i) \times (y_j+z_j) - \lambda_\alpha s_\alpha$	$\frac{1}{2}\rho(\boldsymbol{y}+\boldsymbol{z}).(\boldsymbol{y}+\boldsymbol{z}) - p_\alpha s_\alpha$
$y = \tilde{L}u$	$d_{ij} = u_{j,i}$	$d_{ij} = v_{j,i}$	$\epsilon_{ij} = \frac{1}{2}(v_{i,j}+v_{j,i})$	$v_i = \dfrac{dq_i}{dt}$	$\boldsymbol{y} = \nabla\phi$	$a_{ij}\ddot{q}_j - z_i = \lambda_\alpha A_{\alpha i}$	$\rho\ddot{u}^j - z^j = -(p_\alpha A_\alpha{}^{ij})_{,i}$
$x = \dfrac{\partial Y}{\partial y}$	$s^{ij} = \dfrac{\partial U}{\partial d_{ij}}$	$\dot{s}^{ij} = \dfrac{\partial U}{\partial d_{ij}}$	—	$p_i = \dfrac{\partial L}{\partial v_i}$	$x_i = -\dfrac{\partial p}{\partial y_i}$	$\ddot{q}_i = \dfrac{\partial Y}{\partial y_i}$	$\ddot{u}_j = \dfrac{\partial Y}{\partial y^j}$
$Lx = -\dfrac{\partial Y}{\partial u}$	$-s^{ij}{}_{,i} = g^j$	$-\dot{s}^{ij}{}_{,i} = \dot{g}^j$	$-\sigma_{ij,i} = g_j$	$-\dfrac{dp_i}{dt} = -\dfrac{\partial L}{\partial q_i}$	$-\nabla.\boldsymbol{x} = 0$	$A_{\alpha j}\ddot{q}_j = s_\alpha$	$A_\alpha{}^{ij}\ddot{u}_{j,i} = s_\alpha$
$X(x_i, u_\alpha, \mu, G\mu) = x_i y_i - Y$	$U_c(s^{ij}) + g^j u_j = s^{ij}d_{ij} - U(d_{ij}) + g^j u_j$	$U_c(\dot{s}^{ij}) + \dot{g}^j v_j = \dot{s}^{ij}d_{ij} - U(d_{ij}) + \dot{g}^j v_j$	$\mu f(\sigma_{ij}) + g_j v_j$	$H(p_i, q_i, t) = p_i v_i - L$	$P(\boldsymbol{x}; \alpha, \beta, h, \eta; \nabla h, \nabla\eta) = \boldsymbol{x}.\boldsymbol{y} + p$	$\frac{1}{2}a_{ij}\ddot{q}_i\ddot{q}_j - z_i\ddot{q}_i + \lambda_\alpha s_\alpha$	$\frac{1}{2}\rho\ddot{\boldsymbol{u}}.\ddot{\boldsymbol{u}} - \boldsymbol{z}.\ddot{\boldsymbol{u}} + p_\alpha s_\alpha$
uNx	$u_j n_i s^{ij}$	$v_j n_i \dot{s}^{ij}$	$v_j n_i \sigma_{ij}$	$q_i p_i$	$\phi\boldsymbol{n}.\boldsymbol{x}$	—	$-p_\alpha A_\alpha{}^{ij} n_i \ddot{u}_j$

by Truesdell & Noll (1965, § 88); but the free principle (2.19) itself may be new. A principle valid for a general class of configuration-dependent surface loadings is given by Sewell (1965, equation (43)). Applications of the variational method to solve specific boundary value problems in finite elastostatic strain have been given by Levinson (1965 *a*, *b*) and Huang (1965).

For the present problem it is unreasonable to expect that reciprocal or dual variational principles can *in general* be strengthened to give extremum principles such as (2.42) and (2.44). This is because actual material behaviour makes it clear that requirements of overall convexity such as (2.25) would be too severely restrictive to be insisted upon over the *whole* range of arbitrary reference configurations, in conjunction with boundary conditions of assigned total load or finite displacement on σ. For bifurcation points and turning points of equilibrium load/deflexion paths are known to exist in specific problems, so that the universal uniqueness of finite displacement which could follow from (2.26) (with y as d_{ij}) is ruled out by physical considerations. It follows *a fortiori* that the convexity of the strain energy function $U(d_{ij})$ itself cannot be entertained (Hill 1957 *a*). A detailed study for arbitrary reference configurations of uniqueness theorems ensuing when U is a convex function of variables generating a *subclass* of displacement gradients has been made by Truesdell & Toupin (1963). In many problems there does exist a limited range of (say) loading for which the finite equilibrium displacement *is* unique, and no general approximation method seems yet to have been worked out for estimating this range. As an indication of what may be involved consider a discrete conservative system whose *total* potential energy V is a given function $V(q_i, \lambda)$ of n generalized coordinates q_i whose values can range over a domain D, and of an assignable parameter λ (e.g. load). If there exists a value $\bar{\lambda}$ of λ such that the Hessian $\partial^2 V/\partial q_i \partial q_j$ is positive definite when $\lambda < \bar{\lambda}$ for *all* q_i in D, then the resulting convexity of $V(q_i, \lambda)$ for each such λ shows that there is at most one solution q_i of the equilibrium equations $\partial V/\partial q_i = 0$ when $\lambda < \bar{\lambda}$. The special $q_i (= \bar{q}_i$, say) which cause the Hessian for $\lambda = \bar{\lambda}$ to be semi-definite need not be equilibrium points in general, although they may be so for certain special systems (which require investigation). *When* they are, $\bar{\lambda}$ may be a good estimate of the uniqueness range, but it may not be so in the other cases. (A local investigation of equilibrium paths for discrete systems, and conditions for V to be a local minimum when the above Hessian is semi-definite, are given by Sewell, 1968 *a*, *b* respectively). In the continuum problem the total potential energy involves not only the strain energy but also the potential energy of the external loads. Masur (1954, 1958, 1968) has suggested extremum principles for use in the post-buckling (post-bifurcation) analysis of a class of symmetric structures including statically indeterminate trusses, narrow beams, and plates. In certain cases elastic collapse occurs under a finite 'limit load' which can be bracketed between classes of 'kinetically admissible' and 'statically admissible' load parameters (cf. the special theorems of 'limit analysis' arising from the principle of maximum plastic work and its dual in § 3 (*c*) below).

No significance seems so far to have been attached to (2.34) to (2.36) in the present context (but see §§ 5 and 6 below where internally constrained materials are considered). Unilateral boundary conditions like (2.37) to (2.39) can be proposed for the finite displacement problem, for example in a way analogous to the suggestions (3.1) to (3.3) for the 'instantaneous' boundary value problems discussed in the next example (*b*).

It will be noticed that in this paper we are concerned, among other things, to investigate uniqueness theorems and extremum principles ensuing from essentially a *single* definition of convexity, namely the 'overall convexity' of (2.25). There are, in fact, a number of related but not equivalent definitions of convexity and strong ellipticity. The differing effects of several of these definitions on uniqueness theorems and on other notions of physically reasonable response have been explored in considerable detail for elastic materials by Truesdell & Toupin (1963) and Truesdell & Noll (1965).

(b) *Equilibrium problem of incremental elastic or plastic strain*

This problem is posed in some known configuration which need not be stress-free and which has been arrived at after a strain-history which can have been complicated and which does not require specification for present purposes. This known current configuration is itself employed as the reference configuration, so that the values of nominal stress s^{ij} at this instant (unlike those at later instants) coincide with the given true stress. The body occupies the region τ with boundary σ at this instant. We have to solve the incremental quasi-static problem in which 'velocity' v_j represents displacement-increment and 'stress-rate' $\dot{s}^{ij}$ represents nominal stress-increment (*not* the same as true stress-increment unless all $s^{ij} = 0$). The equations $Lx = s$ are therefore this time $-\dot{s}^{ij}{}_{,i} = \dot{g}^j$, which may be regarded as the first order perturbation equations in at least a formal perturbation scheme (cf. Sewell 1968*a*) expressing continuing equilibrium. Here $\dot{g}^j$ is the assigned nominal body-force rate per unit reference volume. The superposed dot represents differentiation with respect to any 'perturbation parameter' (not real time) whose monotonic mathematical variation generates an ordered sequence of equilibrium states.

The incremental material behaviour is written into equations (2.2) with (2.24), where $U(y)$ this time is an assigned function of velocity gradients $d_{ij} = v_{j,i}$ which is homogeneous of degree two. Then the 'rate equations' $\dot{s}^{ij} = \partial U/\partial v_{j,i}$ are homogeneous of degree one, and Euler's theorem with (2.6) shows that $U_c = U = \frac{1}{2}\dot{s}^{ij}v_{j,i}$ numerically. The simplest case is when $U(v_{j,i})$ is a single quadratic function in a certain class, so that (2.27) applies directly (to U as well as Y)—i.e. the barred quantities on the right in (2.27) are then just the coefficients of this quadratic. Quadratics in this class emerge from stressed and unstressed elastic solids possessing a strain energy (i.e. from the solids in (*a*) above), and also in problems involving hardening plastic solids which are not unloading anywhere. Incremental behaviour of stressed elastic/plastic solids, with the possibility of either loading or unloading in the same material element, may be represented by quadratic functions $U(v_{j,i})$ having discontinuities in their second derivatives. We refer to Hill (1962, 1967) for a further statement of such material properties.

We can again begin with the appropriate form of (2.19), and deduce therefrom the reciprocal variational principles of Hill (1962, § 4). The uniqueness criteria and extremum principles which ensue from the overall convexity of $U(v_{j,i})$ or $U_c(\dot{s}^{ij})$ in the case of bilateral conditions on σ and in τ are also given by Hill (1962, §§ 5 and 6).

Bifurcation of the elastic/plastic column under axial load provides a context in which these variational principles for a composite quadratic U are applicable even when uniqueness is known to fail so that the extremum principles are not valid (see Hill & Sewell 1962). An example of a calculation employing the variational principles to investigate the range of uniqness in a compressed elastic/plastic plate is given by Sewell (1963*b*, 1964). For an elastic solid in its unstressed state the extremum principles reduce to those familiar in classical elasticity (sometimes called Castigliano's principle and its dual).

Although modification of the field equations along the lines of (2.34) to (2.36) has not been given any significance, as in (*a*) above, a similar modification of boundary conditions to include unilateral conditions like (2.37) to (2.39) can here be identified with the presence of *passive constraints*. Let $\dot{t}^j = n_i\dot{s}^{ij}$ denote nominal traction-rate vector on an area element facing in the

$\boldsymbol{n}$-direction. For a simple example suppose that over the whole of σ the following conditions are placed on the normal components of velocity and traction-rate:

$$\boldsymbol{n}\,.\,\dot{\boldsymbol{t}} \leqslant 0, \tag{3.1}$$

$$\boldsymbol{n}\,.\,\boldsymbol{v} \leqslant c, \tag{3.2}$$

$$(\boldsymbol{n}\,.\,\boldsymbol{v}-c)\,\boldsymbol{n}\,.\,\dot{\boldsymbol{t}} = 0. \tag{3.3}$$

Here (3.1) expresses the hypothesis that in the common normal direction the passive constraining surface can only push the body and not pull it, and (3.2) says that the body can leave but not penetrate the constraint which has an assigned normal velocity c (which might vary with position). And (3.3) ensures that if inequality occurs in one of (3.1) and (3.2) then equality must occur in the other (cf. equations (4.11) to (4.13) and (5.18) to (5.20)). Suppose, in addition, that the tangential components of $\boldsymbol{v}$ are assigned on a part σ_v of σ (we could then speak of perfectly rough passive constraints on σ_v), and that the tangential components of $\dot{\boldsymbol{t}}$ are assigned on the remaining part $\sigma_t = \sigma - \sigma_v$ of σ (if these assigned values are zero we could then speak of perfectly smooth passive constraints on σ_t). The affinity of (3.1) to (3.3) with (2.37) to (2.39) is obvious, and we note that the present form of (2.46) is satisfied (integral over σ of the product of the right-hand sides of (3.1) and (3.2) is *zero*).

From the overall convexity of the composite second degree functions $U(v_{j,i})$ and $U_c(\dot{s}^{ij})$ described above it is straightforward to show as in § 2 (vi) that the following complementary extremum principles characterize a solution of the incremental equilibrium problem under the mixed unilateral and bilateral boundary conditions just stated:

$$\begin{aligned}
&\int U_c(\dot{s}^{ij*})\,\mathrm{d}\tau - \int c\boldsymbol{n}\,.\,\dot{\boldsymbol{t}}^*\,\mathrm{d}\sigma - \int[\boldsymbol{n}\wedge(\dot{\boldsymbol{t}}^*\wedge\boldsymbol{n})]\,.\,[\boldsymbol{n}\wedge(\boldsymbol{v}\wedge\boldsymbol{n})]\,\mathrm{d}\sigma_v\\
&\quad\geqslant \int U_c(\dot{s}^{ij})\,\mathrm{d}\tau - \int c\boldsymbol{n}\,.\,\dot{\boldsymbol{t}}\,\mathrm{d}\sigma - \int[\boldsymbol{n}\wedge(\dot{\boldsymbol{t}}\wedge\boldsymbol{n})]\,.\,[\boldsymbol{n}\wedge(\boldsymbol{v}\wedge\boldsymbol{n})]\,\mathrm{d}\sigma_v\\
&\quad= \tfrac{1}{2}\Big[\int \dot{\boldsymbol{g}}\,.\,\boldsymbol{v}\,\mathrm{d}\tau + \int[\boldsymbol{n}\wedge(\dot{\boldsymbol{t}}\wedge\boldsymbol{n})]\,.\,[\boldsymbol{n}\wedge(\boldsymbol{v}\wedge\boldsymbol{n})]\,\mathrm{d}\sigma_t - \int[\boldsymbol{n}\wedge(\dot{\boldsymbol{t}}\wedge\boldsymbol{n})]\,.\,[\boldsymbol{n}\wedge(\boldsymbol{v}\wedge\boldsymbol{n})]\,\mathrm{d}\sigma_v - \int c\boldsymbol{n}\,.\,\dot{\boldsymbol{t}}\,\mathrm{d}\sigma\Big]\\
&\quad= \int[\dot{\boldsymbol{g}}\,.\,\boldsymbol{v} - U(v_{j,i})]\,\mathrm{d}\tau + \int[\boldsymbol{n}\wedge(\dot{\boldsymbol{t}}\wedge\boldsymbol{n})]\,.\,[\boldsymbol{n}\wedge(\boldsymbol{v}\wedge\boldsymbol{n})]\,\mathrm{d}\sigma_t\\
&\quad\geqslant \int[\dot{\boldsymbol{g}}\,.\,\boldsymbol{v}^\dagger - U(v^\dagger_{j,i})]\,\mathrm{d}\tau + \int[\boldsymbol{n}\wedge(\dot{\boldsymbol{t}}\wedge\boldsymbol{n})]\,.\,[\boldsymbol{n}\wedge(\boldsymbol{v}^\dagger\wedge\boldsymbol{n})]\,\mathrm{d}\sigma_t.
\end{aligned} \tag{3.4}$$

Here we have used the same system of designation as in (2.42) and in (2.44). That is, symbols designated by * satisfy the equilibrium equations (2.10) in τ, (3.1) everywhere on σ, and the given tangential conditions on σ_t; but they do not have to be associated with any velocity field via the constitutive rate equations or satisfy (3.2), (3.3) or the tangential conditions on σ_v. Whereas symbols designated by † satisfy (3.2) everywhere on σ and the given tangential conditions on σ_v; but they do not have to be associated with any stress-rate field satisfying (2.10), nor must they satisfy (3.1), (3.3) or the tangential conditions on σ_t. And symbols without either designation have to satisfy *all* the conditions for an actual solution of the boundary value problem. We have repeatedly used the decomposition $\boldsymbol{\lambda} = (\boldsymbol{\lambda}\,.\,\boldsymbol{n})\,\boldsymbol{n} + \boldsymbol{n}\wedge(\boldsymbol{\lambda}\wedge\boldsymbol{n})$ of a vector into its normal and tangential components respectively.

The extremum principles (3.4) are new, and could obviously be extended easily to the possibly more realistic case when the passive constraints apply on only *part* of σ, with assigned $\boldsymbol{n}\,.\,\dot{\boldsymbol{t}}$ or $\boldsymbol{n}\,.\,\boldsymbol{v}$ elsewhere. We repeat that they apply to the incremental behaviour of elastic and elastic/plastic solids. To get the inequalities we have given away not only terms like the integrated right-hand side of (2.27) which are of second order in the 'difference fields', but also $\int \boldsymbol{n}\,.\,\dot{\boldsymbol{t}}^*(\boldsymbol{n}\,.\,\boldsymbol{v}-c)\,\mathrm{d}\sigma \geqslant 0$ to get the minimum principle and $\int \boldsymbol{n}\,.\,\dot{\boldsymbol{t}}(\boldsymbol{n}\,.\,\boldsymbol{v}^\dagger - c)\,\mathrm{d}\sigma \geqslant 0$ to get the maximum principle. In the absence of a guarantee that these extra amounts are not of first order in the difference fields $\Delta\dot{s}^{ij}$ and Δv_j respectively, we are not entitled to suppose that the extrema are stationary.

Uniqueness of the solution characterized by (3.4) is assured by uniqueness theorems of the kind described in §2 (v). For example, the overall convexity of $U(v_{j,i})$ expressed in the form

$$\int \Delta(\partial U/\partial v_{j,i})\, \Delta v_{j,i}\, d\tau > 0 \tag{3.5}$$

ensuing from (2.26), (2.24) and (2.11) for all pairs of distinct velocity-gradient fields is a sufficient criterion for unique $v_{j,i}$—and unique $v_{j,i}$ implies unique v_j apart from a possible constant which is often eliminated by the velocity boundary conditions (cf. Hill, 1959, equation (12)); it is easy to show using (3.1) to (3.3) and the other data that if there could be two distinct solutions of the stated problem the integral in (3.5) would have to be non-positive. Evidently therefore (3.1) to (3.3) could be one example of the sets of boundary conditions envisaged by Hill (1967 equation (3.4)).

(c) *Equilibrium problem of yield-point states*

Attention has already been drawn to the fact that the extremum principles (2.68) may be interpreted as the principle of maximum plastic work and its dual. This interpretation requires the identifications listed in table 1 (c). The continuum in question occupies the region τ with surface σ at the considered instant. This time we use Cartesian spatial coordinates in the current configuration, and σ_{ij} is the symmetric tensor of current true stress, which must satisfy the equilibrium equations $(2.59)_1$ under assigned body forces g_j per unit volume. The incremental displacement components are v_j with associated strain-increments ϵ_{ij}. The statements of (2.10) and (2.11) for this problem, in matrix notation to illustrate (2.8) and (2.9), are therefore

$$L_{kij}\sigma_{ij} = g_k \quad \text{and} \quad \epsilon_{ij} = \tilde{L}_{ijk} v_k \tag{3.6}$$

respectively, where
$$-L_{kij} = \tilde{L}_{ijk} = \frac{1}{2}\left(\delta_{ki}\frac{\partial}{\partial\theta_j} + \delta_{kj}\frac{\partial}{\partial\theta_i}\right). \tag{3.7}$$

Here $(\theta_1, \theta_2, \theta_3)$ are the spatial Cartesian coordinates, and we have used the symmetry of σ_{ij} in $(3.6)_1$. By contrast it was only the *second* part of the expression in (3.7) which was required in the above problems (*a*) and (*b*) where the nominal stress s^{ij} and stress-rate $\dot{s}^{ij}$ were in general unsymmetric. The function f in (2.58) is an assigned function of stress and perhaps also position which characterizes the local 'strength' of the body, for example in the generalized sense described by Hill (1966*a*). In particular $f(\sigma_{ij})$ can be the yield function of a rigid/plastic body, which may be hardening, non-hardening or softening. When $f(\sigma_{ij})$ is symmetrized with respect to σ_{ij} and identified with the plastic potential the associated 'flow rule' takes the form $(2.61)_1$, where μ is a proportionality factor which does not require further specification here. The relation between the existence or otherwise of local yield point states and of locally deforming modes is then described by figure 3 (*c*) or (2.69) to (2.71).

The general formalism may be regarded as beginning with the function

$$X(\sigma_{ij}, v_j, \mu) = \mu f(\sigma_{ij}) + g_j v_j \tag{3.8}$$

inserted into (2.4), thus defining the y_i as the strain-increments ϵ_{ij}, these latter being related to the displacement-increments by (2.11) in the form $(3.6)_2$ (which for brevity we have employed at the outset in the free variational principle this time). We do not need to suppose that ensuing equations $\epsilon_{ij} = \partial X/\partial\sigma_{ij}$ are invertible, so that a complementary function Y independent of stress is never defined. As implied after (2.63) the uniqueness theorem and extremum principles are based on the alternative statements (2.29) and (2.32) of the convexity of X with respect to the six independent components σ_{ij} at each fixed μ. Actually such convexity is stricter than that usually found, at least in metals, where the yield function $f(\sigma_{ij})$ is commonly regarded as independent of the hydrostatic component of stress so that $>$ is replaced by $\geqslant$ in (2.64) and (2.65); but this qualification does not necessarily apply in the generalized problem stated by Hill (1966*a*).

Under the boundary conditions of assigned load on σ_x and assigned displacement on σ_u ($(2.59)_2$ and $(2.61)_2$ respectively) the uniqueness theorem stemming from (2.72) ensures that the stress distribution, or at least its deviatoric part, is unique in the deformable zone. This is shown by Hill (1951), who also explains how the upper and lower bounding theorems of the so-called 'plastic limit analysis' are special cases of his earlier statements of the extremum principles (2.68). The non-analytic character of these extremum principles makes it natural to ask whether established techniques of 'mathematical programming' can be exploited in their applications, and Charnes, Lemke & Zienkiewicz (1959) have expressed the static and kinematic plastic collapse principles for frames as dual linear programming problems. We note that the uniqueness theorem, involving as it does a contradiction of a convexity property with respect to stress only, can say nothing about the uniqueness or otherwise of μ or of the deformation modes; more data are required for an examination of this question (see Hill 1957*b*, *c*), which involves a boundary value problem like that in example (*b*) above.

Collins (1968) has drawn attention to another type of passive constraint, which may be expressed in terms of the tangential components of the surface traction $t_j = n_i \sigma_{ij}$ acting on the body and of its velocity $\boldsymbol{v}$, in a form reminiscent of the passive normal constraints (3.1) to (3.3). For example, the maximum shear stress $\hat{\boldsymbol{t}}.[\boldsymbol{n}\wedge(\boldsymbol{t}\wedge\boldsymbol{n})] = t$ (say, where $\hat{\boldsymbol{t}}$ is a unit vector in the unassigned direction of maximum shear) exerted by the constraining surface may take any non-negative value not exceeding an envisaged assigned local yield stress $t_Y > 0$, but relative motion between the constraint and the body in question, in the opposite direction to the shear, can only take place if $t = t_Y$. When the passive constraint is fixed and rigid let $-\hat{\boldsymbol{t}}.[\boldsymbol{n}\wedge(\boldsymbol{v}\wedge\boldsymbol{n})] = v$ (say) denote the tangential speed of the body. Then analytically we have

$$0 \leqslant t \leqslant t_Y, \tag{3.9}$$

$$v \geqslant 0, \tag{3.10}$$

$$(t_Y - t)\, v = 0, \tag{3.11}$$

(cf. the 'stiction' of Hill 1963*a*, equation (2.6)). The affinity both of these and of the earlier passive surface constraints to figure 2(*c*) is obvious. Under suitable boundary conditions on the normal traction and velocity (such as assigned values) on this part of the surface, and on the traction or velocity elsewhere on the boundary, the uniqueness theorem is again valid, because the difference of any two solution pairs contributes to the surface integral arising in the integrated left-hand side of (2.72) an amount $[\boldsymbol{n}\wedge(\Delta\boldsymbol{v}\wedge\boldsymbol{n})].[\boldsymbol{n}\wedge(\Delta\boldsymbol{t}\wedge\boldsymbol{n})]$ which is non-positive by (3.9) to (3.11). As we by now anticipate from the analytical structure, the maximum plastic work principle and its dual in (2.68) can also both be generalized to include these passive constraints.

The maximum work principle maximizes $\int \boldsymbol{h}.\boldsymbol{t}\,\mathrm{d}\sigma_v$ when $\boldsymbol{v} = \boldsymbol{h}$ on σ_v, and $\int h\boldsymbol{n}.\boldsymbol{t}\,\mathrm{d}\sigma_v$ when $\boldsymbol{n}.\boldsymbol{v} = h$ and (3.9) to (3.11) apply on σ_v. Collins (1968) interprets these results as maximizing a 'total load' by giving special interpretations to the assigned functions of position $\boldsymbol{h}$ and h.

(d) *Classical dynamical problem*

For perspective we mention this problem here, because up to a point it serves as the historical prototype for the general theory of §2. In table 1 (d) there are n generalized coordinates q_i and the real time t, which act as passive variables in the Legendre transformation relating the assigned Lagrangian function $L(v_i, q_i, t)$ to the Hamiltonian $H(p_i, q_i, t)$. Here the n p_i are the momenta, and (2.8) and (2.9) are the equations of motion of the holonomic system, and the equations relating the v_i to the generalized velocities, respectively. The operators in (2.8) and (2.9) are defined by $\tilde{L}_{i\alpha} = \delta_{i\alpha}\mathrm{d}/\mathrm{d}t$ and $L_{\alpha i} = -\delta_{\alpha i}\mathrm{d}/\mathrm{d}t$. The 'boundary' σ is now just the two end-points of a time interval τ between two t-values, t being the only independent variable. The free variational principle (2.18) is just an extended form of Hamilton's principle with time integrand $p_i v_i + q_i \dot{p}_i - L(v_i, q_i, t)$ in which *all* of the p_i, v_i and q_i are treated as unrelated variables at the outset. Comparatively little is said in the literature about the uniqueness theorems, extremum principles, etc., which further restrictions would produce.

(e) *Equations of inviscid fluid mechanics*

When we consider the class of all materials for which the stress is derivable from a strain energy, and in particular (hyper-)elastic solids and perfect fluids, it is clear that this strain energy is going to appear in some form in any variational principle which leads to the balance equations of linear momentum, and that such principles will be related to Hamilton's principle (see, for example, Truesdell & Toupin, 1960, §§232A, 236A). This will give an indication of the function Y to be employed in (2.16), but the particular purpose in hand will also require a decision whether the region τ should best be chosen as the union of an (arbitrary) time interval with, for example, some fixed (perhaps previous) reference volume or with the current space occupied by the material. In Hamilton's principle itself, and for the special (equilibrium) class of solutions sought in §3 (a) the reference volume is appropriate. For practical purposes in problems involving motion it is often the current spatial coordinates which are more convenient to use and, in addition, we may then wish the variational principle to yield other equations such as the equations of continuity and energy, in addition to the equations of linear momentum. The strain energy may then be identified with the specific internal energy or free energy in certain well-defined circumstances (see, for example, Truesdell & Toupin 1960, §256A). In such problems an approach too closely tied to Hamilton's principle can become very cumbersome, as several investigators have remarked (notably Seliger & Whitham 1968), and it is by no means obvious how the integrand in (2.16) should be completed in order that (2.18) should furnish the general equations.

Our purpose here is not so much to seek any canonical expression of the *general* equations, but rather to point out another special class of problems (in addition to those of §3 (a)) in which the formalism of §2, and in particular the existence of extremum principles strengthening the variational principles, can be verified. This class of problems concerns the steady flow of perfect fluids.

As the starting point we suppose that the internal energy ϵ per unit mass is given as a function

$$\epsilon = \epsilon(\eta, \rho) \tag{3.12}$$

of density $\rho > 0$ and entropy η per unit mass. Definitions of thermodynamic pressure $p \equiv -\partial\epsilon/\partial(1/\rho)$ and of velocity of sound c given by $c^2 \equiv \partial p/\partial \rho$ are then introduced as functions of η and ρ. Some universally valid functional inversions then lead to the following results for a compressible fluid. We can express p as an assigned function

$$p = p(\eta, \chi) \tag{3.13}$$

such that
$$\frac{\partial p}{\partial \chi} = \rho, \quad \frac{\partial p}{\partial \eta} = -\rho\theta, \quad \frac{\partial^2 p}{\partial \chi^2} = \frac{\partial \rho}{\partial \chi} = \frac{\rho}{c^2}. \tag{3.14}$$

Here $\theta > 0$ is absolute temperature, and χ having the value of $\epsilon + p/\rho$ is the enthalpy of unit mass. The independent variables in the differentiations in (3.14) are η and χ only, so that (3.14) may be regarded as expressing ρ, θ and c^2 as assigned functions of η and χ. For simplicity we shall consider below only the homogeneous case, in which the function (3.12) is the same function for each fluid particle.

We next choose τ in (2.12), (2.16) *et seq.* to be the region of space occupied by the steady motion under discussion, and we investigate the possibility that a suitable function Y shall have the form

$$Y(\rho, \eta, K) = \rho[\epsilon(\eta, \rho) - K]. \tag{3.15}$$

Here K is to be regarded as a function of η and of certain other variables which we have to choose in a way which will allow us to get the equations of continuity, energy and motion from (2.18). Since more groups of equations are sought than in the other problems illustrated in table 1, we expect that a form of type (2.54) might result, in which several passive variables μ and perhaps their gradients $G\mu$ will appear; the corresponding extra groups of equations will then emerge from natural conditions like (2.55). The choice (3.15) reflects the indication from Hamilton's principle that the mass-integrand will contain ϵ for motions with no thermodynamic constraints, together with 'something else' K. If ρ does not appear explicitly in K or in the x, y and u of (2.16), then as shown in Sewell (1963*a*, equation (33)) one of the natural conditions of (2.18) is

$$\delta\rho: \quad Y = -p. \tag{3.16}$$

Thus the value of (3.15) wherever (3.16) holds is that of minus the thermodynamic pressure, and the value of $K = \epsilon + p/\rho$ will then be that of χ itself. For an incompressible fluid we may begin with (3.15) but with ρ regarded as a fixed parameter so that $\delta\rho \equiv 0$ and (3.16) does not arise.

Let us suppose that (3.16) and the consequent $K = \chi$ are assumed at the outset. The starting-point now is therefore the function

$$Y(\eta, \chi) = -p(\eta, \chi) \tag{3.17}$$

defined by (3.13). Hamilton's principle also suggests that an appropriate mass integrand will contain $\frac{1}{2}v_i^2$, where $\boldsymbol{v}$ is fluid velocity, so in view of (3.15) and $K = \chi$ we write

$$\chi = -\tfrac{1}{2}v_i^2 + h \tag{3.18}$$

where h is a new variable.

It is shown by Sewell (1963*a*) that a Legendre dual transformation can be constructed beginning with (3.17) and (3.18), treating the (Cartesian) components v_i of $\boldsymbol{v}$ as active variables and η and h as passive. The values of the derivatives of (3.17) with respect to the active variables are

$$-\frac{\partial p}{\partial v_i} = \rho v_i, \quad -\frac{\partial^2 p}{\partial v_i \partial v_j} = \frac{\rho}{c^2}(c^2\delta_{ij} - v_i v_j). \tag{3.19}$$

The transformation is reversible except at transonic points where $v = c$, and the function dual to $-p$ has the value of $p + \rho v^2$. This leads to a free variational principle like (2.18) from which emerges a 'Clebsch representation' in the form

$$\boldsymbol{v} = \nabla\phi + \alpha\nabla h + \beta\nabla\eta, \tag{3.20}$$

together with the equation of continuity, the equation of energy, the Bernoulli integral that h be constant on streamlines, and the Crocco–Vaszonyi form of the equations of motion without body forces (on the supposition that p is also the mechanical pressure). It is of interest that the statement of the free variational principle in this case of steady motion can be reasonably perceived on the basis of the expected forms of the desired resulting equations, and it does not need hindsight derived from other sources (the same cannot be said of the corresponding principle for unsteady motion, as a study of the literature makes plain). In (3.20) ϕ, α and β are functions of position which are introduced as Lagrangian multipliers in the free principle, and ∇ is the spatial gradient operator. Ensuing reciprocal principles are also proved by Sewell (1963 *a*) under suitable boundary conditions permitting a form of (2.19), and for homentropic ($\eta \equiv$ constant), homenergic ($h \equiv$ constant) steady flows a uniqueness theorem and dual extremum principles are proved which are valid when the flow is subsonic in an overall sense. These extremum principles rest on the facts that $-p(v_i)$ and its dual are convex functions when $v < c$, as may be inferred from $(3.19)_2$ in (2.28). The principles are always valid in the incompressible case, as may be verified by beginning with (3.15) with fixed ρ and with $K = h - \frac{1}{2}v_i^2$.

The precise correspondence of the analysis in the paper just mentioned to that of § 2 requires the modifications implied in column (*e*) of table 1. These modifications are purely formal except that the Clebsch representation has to be assumed (in effect), instead of being allowed to emerge naturally as it does in the paper. For instead of (3.18) we take

$$\chi = h - \tfrac{1}{2}(\boldsymbol{y} + \alpha\nabla h + \beta\nabla\eta)^2 \tag{3.21}$$

in (3.17), so that the resulting function Y is of type (2.54), namely

$$Y(\boldsymbol{y}, h, \eta, \alpha, \beta; \nabla h, \nabla\eta) = -p(\eta, h - \tfrac{1}{2}(\boldsymbol{y} + \alpha\nabla h + \beta\nabla\eta)^2). \tag{3.22}$$

The value of $\boldsymbol{x}$ in the table 1 (*e*) is just the 'mass flow' $\rho\boldsymbol{v}$ (see $(3.19)_1$), $Lx = 0$ represents the continuity equation, and $\tilde{L}u$ is $\nabla\phi$ with ϕ not appearing explicitly in Y so that $s \equiv \partial Y/\phi = 0$. In these terms the free principle (14) of Sewell (1963 *a*) exemplifies (2.18).

It is possible to generalize the Legendre transformation and free variational principle to cover the case of unsteady flows also. For example, it is clear from Seliger & Whitham (1968, equation (40)) that a free variational principle may be obtained by replacing (3.18) and (3.20) by

$$\chi = \mathscr{H}(\alpha, h, t) - \tfrac{1}{2}(\nabla\phi + \alpha\nabla h + \beta\nabla\eta)^2 - (\partial\phi/\partial t + \alpha\partial h/\partial t + \beta\,\partial\eta/\partial t), \tag{3.23}$$

where $\mathscr{H}$ is a known function of the indicated variables having the properties

$$\frac{\mathrm{D}h}{\mathrm{D}t} = \frac{\partial\mathscr{H}}{\partial\alpha}, \quad \frac{\mathrm{D}\alpha}{\mathrm{D}t} = -\frac{\partial\mathscr{H}}{\partial h}, \tag{3.24}$$

with $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \boldsymbol{v}.\nabla$; and by taking τ to be the union of an (arbitrary) time interval with the (varying) spatial region occupied by the fluid during that time. For the verification that the equations of continuity, energy and motion ensue we refer to Seliger & Whitham (1968), who also treat the general motion of an elastic solid from the same viewpoint. The detailed analysis of unsteady motions, and in particular the question whether any of the variational principles can be strengthened to provide extremum principles, would require too extensive a digression to be included here. The reader may refer to the review article by Serrin (1959) for a statement of the position which studies of variational principles in fluid mechanics had reached up to that time.

Seliger & Whitham were apparently unaware of the author's 1963 paper, and they do not mention the central roles of the Legendre transformation and of convex functions. In fact they are not concerned with extremum principles, most of their emphasis being on variational principles for perfect fluids (and to a lesser extent elastic solids). What seems to be their search for reciprocal principles (cf. §2 (iii) here) ends in a rather inconclusive examination of a theorem of Pfaff. Nevertheless, Whitham (1967) has applied variational methods to the approximate solution of problems in water waves (see also Luke (1967)).

It is to be recognized that unless a variational principle can be strengthened to produce an extremum principle, there is no objective assurance that an approximate solution based upon it will be better than that obtainable from some other 'averaging' principle or *ad hoc* approximation. Examples of such situations in the theory of metal working processes and in plastic column bending, in which an averaging principle based on the virtual work principle is compared with a variational principle, are discussed by Hill (1963 *a*) and Hill & Sewell (1962) respectively.

The differences between the principles in this subsection and in §3 (*a*) are noteworthy. For example, (2.10) has here the role of the continuity equation and involves spatial coordinates, whereas (2.10) has there the role of equilibrium equations and involves material coordinates. It would be interesting to know the results of a full examination of the situation in which extremum principles can be proved for materials having a strain energy, since the special examples brought together here and in §3 (*a*) display little direct connexion apart from the common mode of approach given in §2.

We have hardly mentioned boundary conditions, which might be included by augmentation as in the transition from (2.18) to (2.22). However, to illustrate the range of problems in which variational methods have actually been used in fluid mechanics, we cite the applications by Moiseev & Petrov (1966) for internal flow problems, by Fiszdon (1962) and Lush & Cherry (1956) for external flow problems, by Krajewski (1963) to the flow through turbomachinery, and the work of Luke and Whitham mentioned above on problems involving a free surface. That such methods may nevertheless not have been fully exploited is suggested by the fact that the necessary general theorems have not yet been fully stated or explored, as this subsection shows.

(*f*) *Finite-dimensional initial motion problem*

This problem is explained in §4, and summarized in table 1 (*f*).

(*g*) *Initial motion problem in a continuum*

This problem is explained in §5, summarized in table 1 (*g*), and illustrated for the incompressible Newtonian viscous fluid in §6. The practical viability of the new principles proved in §§5 and 6 remains to be tested.

(*h*) *Other examples*

Other illustrations of problems having the structure outlined in §2 are readily found. Hill (1956) indicates a number of these, with emphasis on the equilibrium problem (§3 (*a*) or §3 (*b*)) for a variety of material properties. The extremum principles which bound measures of drag in slow flow of the Newtonian viscous fluid (Hill & Power 1956) provide one example. The convex function there is the dissipation, which is a quadratic function of strain-rate formally similar to the strain-energy of the classical isotropic linear elastic solid. Recent generalizations for the incremental equilibrium problem cover material properties which are inhomogeneous in a discontinuous fashion, as in multiphase solids (Hill 1963 *b*, *c*) and aggregates of crystals (Hill 1966 *b*).

Gladwell & Zimmermann (1966) have produced reciprocal variational principles for a class of acoustic and structural linearized free vibration problems, and Gladwell (1965) has suggested that the finite element method may then be employed to generate approximate solutions. The time-dependence is assumed to be exponential (harmonic) and removed at the outset, so that the variational principles characterize a time-independent problem (being to that extent like the problems (*a*) and (*b*) or (*f*) and (*g*) here). It is a simple matter to express their equations in the form of § 2. It may then be shown, for example, in the coupled problem of air vibrating in a rigid cavity closed on one side by an elastic membrane, that all the governing equations (and in particular their reciprocal principles) are obtainable from a free variational principle similar to (2.19) here. They say nothing about uniqueness theorems or extremum principles.

The isotropic diffusion equation $\nabla^2\phi + a^2\phi = b$ can be put into the form of (2.47) and (2.48) by introducing the subsidiary variable $\boldsymbol{u} = -\nabla\phi$. One may then apply the theory of § 2 (vii) to the function

$$X(\phi, \boldsymbol{u}) = \tfrac{1}{2}\boldsymbol{u}^2 - \tfrac{1}{2}a^2\phi^2 + b\phi, \tag{3.25}$$

which has the requisite saddle shape mentioned.

Clearly it is not difficult to find cases where the full potential of § 2 has not been realized. Other contexts requiring investigation are magnetofluiddynamics and viscous flow with significant inertia terms, in both of which fields variational principles have been employed (see, for example, Cap & Müller (1967), and Schecter (1966), respectively).

The optimum design of structures is an active field which is a little removed from the spirit of this paper, but it seems worth mentioning briefly one or two points of contact. The variational principles and extremum principles of §§ 3 (*a*), (*b*) and (*c*) carry the implicit assumption that the layout or distribution τ of the material is specified, and in that context are therefore fundamental tools of structural *analysis*. Structural *design* is the problem of varying the layout to suit some specific optimizing purpose, and Prager & Taylor (1968) illustrate how such variational principles may be extended to structural design for minimum weight. Optimum design problems often reduce to mathematical programming problems which may be linear (see, for example Hemp & Chan 1966), or non-linear (see, for example Chan (1968) or Pope (1968)), but the associated unilateral constraints can arise from considerations other than those mentioned in this paper.

4. Finite-dimensional initial motion problem

To illustrate the theoretical structure of § 2 in the purely algebraic case we generalize some results recently obtained by Moreau (1966*a*). The problem concerns a mechanical system whose typical particle has a position vector $\boldsymbol{r}$ which may be expressed with respect to a fixed background reference frame as a given function

$$\boldsymbol{r} = \boldsymbol{r}(q_i, t) \tag{4.1}$$

of $n+m$ generalized coordinates q_i and the time t. The system is subject to m kinematical restrictions (not necessarily holonomic) of the form

$$\left.\begin{aligned} A_{\alpha i}\dot{q}_i + B_\alpha &= 0 \quad (\alpha = 1, \ldots, m),\\ A_{\alpha i} &= A_{\alpha i}(q_j, t) \quad (i, j = 1, \ldots, n+m),\\ B_\alpha &= B_\alpha(q_j, t), \end{aligned}\right\} \tag{4.2}$$

where

are assigned functions of the indicated variables. The $n+m$ equations of motion are

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \frac{\partial T}{\partial q_j} - Q_j - \lambda_\alpha A_{\alpha j} = 0, \tag{4.3}$$

if the constraints are workless in the usual sense implied by the derivation from D'Alembert's principle. Here the λ_α are m Lagrangian multipliers associated with the constraints, the generalized forces $Q_j = Q_j(q_i, \dot{q}_i, t)$ are assigned functions of the indicated variables, and the kinetic energy $T = T(q_i, \dot{q}_i, t)$ has the form

$$T = \tfrac{1}{2}\dot{q}_i\dot{q}_j a_{ij}(q_k, t) + \dot{q}_i b_i(q_k, t) + c(q_k, t) \tag{4.4}$$

which (4.1) implies. We assume that the $(n+m)\times(n+m)$ matrix

$$a_{ij} \text{ is positive definite.} \tag{4.5}$$

Suppose now that, instead of seeking the full solution $q_i = q_i(t)$, $\lambda_\alpha = \lambda_\alpha(t)$ of (4.2) and (4.3) for all time t, we wish to solve the following initial value problem. At some given instant the values of the generalized displacements q_i and velocities $\dot{q}_i$ are supposed known (and satisfy (4.2)). What are the values at that instant of the generalized accelerations $\ddot{q}_i$ and of the λ_α (or at least of the reactions $\lambda_\alpha A_{\alpha j}$)? The $n+2m$ governing equations of *this* problem have the form

$$A_{\alpha j}\ddot{q}_j = s_\alpha, \tag{4.6}$$

$$a_{ij}\ddot{q}_j - z_i = \lambda_\alpha A_{\alpha i}, \tag{4.7}$$

from (4.2) (differentiated) and (4.3) respectively. Here the m terms s_α and the $n+m$ terms z_i are known in terms of the given data, being assigned functions of the q_i, $\dot{q}_i$, t. Evidently (4.6) and (4.7) are examples of (2.8) and (2.9) respectively, with (2.15) and (2.24), and the entire theory of §2 may be applied to this problem by adopting the correspondences listed in column (f) of table 1.

Since (2.28) is satisfied by the assumption (4.5), the statement of the ensuing general theorems for this problem may be written down at sight by inserting these correspondences into §2, and dropping the τ-integrations and all reference to boundary terms. Therefore we shall not pause to list the detailed results. However, some characteristic features of this problem deserve comment.

(i) X and Y here are single quadratics so that the uniqueness and extremum principles can be proved by elementary inequalities without a direct reference to convexity—but such an approach obscures the fact that the general procedure of §2 applies just as well to non-quadratic convex functions. For example, in the steady incompressible case of §3 (e) quadratic convex functions appear in the proof of Kelvin's theorem of minimum kinetic energy, but in the steady compressible case the uniqueness and extremum principles depend on non-quadratic functions which are convex when the flow is subsonic (Sewell 1963*a*).

(ii) The uniqueness theorem guarantees unique y_i and therefore unique acceleration $\ddot{q}_i$, but the associated λ_α are not necessarily unique until we make the extra assumption that the constraints are independent in the sense that at least one of the $m\times m$ minors of the $m\times(n+m)$ matrix $A_{\alpha i}$ is non-zero (thus permitting m of the linear equations (4.7) to be solved for the λ_α in terms of the unique y_i). Also, the *reductio ad absurdam* proof of the uniqueness of $\ddot{q}_i$ and y_i does not, of itself, guarantee the existence of a solution—the same remark applies to problems involving linear differential operators. In this problem, however, we may regard (4.6) and (4.7) as linear equations whose matrix expression is

$$\begin{bmatrix} \boldsymbol{a} & \boldsymbol{A} \\ \boldsymbol{A}^T & \boldsymbol{O}_m \end{bmatrix}\begin{bmatrix} \ddot{\boldsymbol{q}} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \boldsymbol{z} \\ \boldsymbol{s} \end{bmatrix} \tag{4.9}$$

in the obvious notation. It may be shown via the Laplace expansion (used twice, down the last m rows and columns), and the properties of compound matrices, that the determinant of coefficients in (4.9) is non-zero when (4.5) holds and at least one of the $m\times m$ minors of $\boldsymbol{A}$ is non-zero. Hence a unique solution of (4.9) for $[\ddot{\boldsymbol{q}}\,\boldsymbol{\lambda}]^T$ exists because it can be found by the direct method.

(iii) The extremum principle (2.42) for the present case with bilateral conditions is of course often called Gauss's principle of least constraint, since $U_c(x^*) = \frac{1}{2}a_{ij}\ddot{q}_i^*\ddot{q}_j^* - z_i\ddot{q}_i^*$ is minimized among the solutions of (4.6). The lower bound in the dual principle (2.44) is a non-homogeneous quadratic

$$\lambda_\alpha^\dagger s_\alpha - \tfrac{1}{2}a_{ij}^{-1}(\lambda_\alpha^\dagger A_{\alpha i} + z_i)(\lambda_\beta^\dagger A_{\beta j} + z_j) \tag{4.10}$$

in the $\lambda_\alpha^\dagger$ (which are unrestricted in the bilateral case so far considered), and may therefore be improved by maximizing with respect to the $\lambda_\alpha^\dagger$.

(iv) The form of (2.34) to (2.36) in the present case is

$$A_{\alpha j}\ddot{q}_j \geqslant s_\alpha, \tag{4.11}$$

$$\lambda_\alpha \geqslant 0, \tag{4.12}$$

$$\lambda_\alpha[A_{\alpha j}\ddot{q}_j - s_\alpha] = 0, \tag{4.13}$$

and a graphical representation like that of figure 1 (*b*) could be given. When these are used in place of (4.6) the $\ddot{q}_i$ are still found to be unique, and the statement of extremum principles are unchanged (except that in the lower bound (4.10) the $\lambda_\alpha^\dagger$ must now satisfy $\lambda_\alpha^\dagger \geqslant 0$). However, the analytic character of the extrema no longer holds in general.

(v) The foregoing theory applies in particular when the system is subject to holonomic constraints which for all time may be either

$$(a) \quad \text{bilateral} \quad f_\alpha(q_i, t) = 0, \tag{4.14}$$

or

$$(b) \quad \text{unilateral} \quad f_\alpha(q_i, t) \geqslant 0. \tag{4.15}$$

In either case we have

$$A_{\alpha i} = \frac{\partial f_\alpha}{\partial q_i}, \quad B_\alpha = \frac{\partial f_\alpha}{\partial t}, \tag{4.16}$$

in (4.2) and subsequently. At the instant of setting the initial value problem the given data are assumed consistent with (4.14) and (4.2) (i.e. $\mathrm{d}f_\alpha/\mathrm{d}t = 0$) then, in both bilateral and unilateral cases. But (4.15) leads to (4.11) (i.e. $\mathrm{d}^2 f_\alpha/\mathrm{d}t^2 \geqslant 0$) at the given instant, while (4.14) leads to (4.6) then. Equations (4.12) associate a conventional sign with the reactions associated with the surfaces of constraint, while (4.13) ensures that if the system moves away from any such passive boundary (so that >0 *then* applies in (4.11)) the associated λ_α vanishes. Moreau (1966*a*) has examined the initial value problem for the holonomic unilateral case (4.15) and he asserts that the uniqueness theorem and (generalized) Gauss principle may be derived from Kuhn and Tucker's theory of multipliers in nonlinear programming. The characterization of a passive boundary in the present problem may be contrasted with that described in the 'quasi-static' problem of § 3 (*b*).

5. Initial motion problem in a continuum

We are concerned here with uniqueness and extremum principles associated with an initial value problem for certain continuous media, which is the analogue of the finite-dimensional problem discussed in § 4. Material coordinates $\theta^i (i = 1, 2, 3)$ are assigned to the particles of the medium, and in any particular configuration these θ^i may be expressed in terms of a background frame of fixed rectangular Cartesian coordinate planes. The inverse of such expressions for some chosen 'reference configuration' may be summarized by writing

$$\boldsymbol{r} = \boldsymbol{r}(\theta^i), \tag{5.1}$$

where $\boldsymbol{r}(\theta^i)$ is the position vector, referred to the Cartesian frame, of the typical particle in the reference configuration. Consequent base vectors for the reference configuration will be denoted by $\boldsymbol{g}_i = \partial \boldsymbol{r}/\partial\theta^i$ and $\boldsymbol{g}^i$, and the medium is supposed to occupy a volume τ with area σ then. Denote the displacement from the reference configuration to some 'current' configuration after (real) time t by

$$\boldsymbol{u} = \boldsymbol{u}(\theta^i, t) = u^i \boldsymbol{g}_i, \tag{5.2}$$

so that $\boldsymbol{u}(\theta^i, 0) = 0$ and the new position vector is $\boldsymbol{u}+\boldsymbol{r} = \boldsymbol{R}(\theta^i, t)$ (say).

It is possible to envisage a very general class of non-holonomic kinematical constraints described by nonlinear assigned functions of t, $\boldsymbol{r}$, $\boldsymbol{u}$, $\dot{\boldsymbol{u}}$ (= particle velocity $\partial\boldsymbol{u}/\partial t|_{\theta^i=\text{const.}}$) and their derivatives with respect to the θ^i, as in Truesdell & Toupin (1960, equation (237.10)). To illustrate our theory we shall consider for simplicity only those constraints which have the property

$$A_\alpha^{ij}\dot{u}_{j,i} + B_\alpha = 0, \quad (\alpha = 1, \ldots, m) \tag{5.3}$$

in which the A_α^{ij} and B_α are assigned functions of t and of $\boldsymbol{u}$ and its θ^i-derivatives (the subscripts following a comma indicate covariant differentiation using the reference base vectors: e.g. $\dot{u}_{j,i} = \boldsymbol{g}_j \cdot \partial\dot{\boldsymbol{u}}/\partial\theta^i$). Conditions (5.3) may arise, for example, as the time-differentiated form of holonomic constraints such as

$$C_\alpha\left(\boldsymbol{r}, \frac{\partial \boldsymbol{r}}{\partial\theta^i}, \frac{\partial \boldsymbol{u}}{\partial\theta^i}, t\right) = 0, \tag{5.4}$$

and in such a case

$$A_\alpha^{ij} = \frac{\partial C_\alpha}{\partial u_{j,i}}, \quad B_\alpha = \frac{\partial C_\alpha}{\partial t}. \tag{5.5}$$

On the other hand, (5.3) may be assigned directly as non-holonomic constraints and regarded as a special case of the type

$$C_\alpha\left(\boldsymbol{r}, \frac{\partial \boldsymbol{r}}{\partial\theta^i}, \frac{\partial \boldsymbol{u}}{\partial\theta^i}, \frac{\partial \dot{\boldsymbol{u}}}{\partial\theta^i}, t\right) = 0, \tag{5.6}$$

in which the functions C_α are linear in the velocity gradients but not integrable, so that they could not be derivable from (5.4).

Now let $\rho = \rho(\theta^i)$ be the density in the reference configuration, and let the current assigned body-force per unit mass be $\boldsymbol{b} = \boldsymbol{b}(\theta^i, t) = b^i \boldsymbol{g}_i$, so that the b^i are 'nominal' components—i.e. components referred to the base vectors as they were in the reference configuration. Likewise let the current stress vector per unit reference area on an area element which had unit outward normal $\boldsymbol{n} = n_i \boldsymbol{g}^i$ in the reference configuration be $\boldsymbol{t} = t^i \boldsymbol{g}_i$. Also let

$$\boldsymbol{S} = s^{ij}\boldsymbol{g}_i\boldsymbol{g}_j \tag{5.7}$$

be the (unsymmetric) tensor of current nominal stress (Hill 1957*a*). Then Cauchy's formula takes the form

$$t^j = n_i s^{ij} \quad \text{on} \quad \sigma \tag{5.8}$$

and the differential equations expressing the current balance of linear momentum are

$$s^{ij}{}_{,i} + \rho b^j = \rho \ddot{u}^j \quad \text{in} \quad \tau. \tag{5.9}$$

We consider any material in which s^{ij} has the form

$$s^{ij} = r^{ij} - p_\alpha A_\alpha^{ij} \quad (\alpha \text{ summed}), \tag{5.10}$$

in which r^{ij} is calculable with the aid of constitutive equations, and the p_α are m multipliers unknown in advance and sought from the boundary value problem as functions $p_\alpha(\theta^i, t)$. (These multipliers will need to be tensors $\boldsymbol{p}_\alpha$ if the original equations of constraint are tensorial, i.e. $\boldsymbol{C}_\alpha = 0$, instead of scalar functions as in (5.4) and (5.6).) The 'virtual work' of the stresses for any differentiable vector field $\boldsymbol{\eta} = \eta_j \boldsymbol{g}^j$ is $s^{ij}\eta_{j,i}$. Now for those special $\boldsymbol{\eta}$ which are interpretable as the difference between any two velocity fields which satisfy the constraints (5.3) everywhere at a given instant we have

$$A_\alpha^{\ ij}\eta_{j,i} = 0. \tag{5.11}$$

The virtual work of the stresses for *such* $\boldsymbol{\eta}$ therefore contains no contribution from the 'reactions' $p_\alpha A_\alpha^{\ ij}$, by (5.10) and (5.11), and in this sense the internal constraints (5.3) in a material of type (5.10) are said to be smooth and the associated reactions workless.

We now define an initial motion problem in the current configuration. The current configuration and velocity distribution are supposed known, together with any other properties of the strain history entering r^{ij}, so that this part r^{ij} of the stress is regarded as assigned at the considered instant. These given data are also supposed to satisfy the constraints (5.3) (and (5.4) where defined) at the current instant. It follows from (5.3) (differentiated) and (5.9) with (5.10) that the acceleration at the given instant must satisfy

$$A_\alpha^{\ ij}\ddot{u}_{j,i} = s_\alpha \tag{5.12}$$

and

$$\rho\ddot{u}^j - z^j = -(p_\alpha A_\alpha^{\ ij})_{,i}, \tag{5.13}$$

where

$$s_\alpha = -(\dot{B}_\alpha + \dot{A}_\alpha^{\ ij}\dot{u}_{j,i}) \quad \text{and} \quad z^j = r^{ij}_{\ \ ,i} + \rho b^j \tag{5.14}$$

are known from the given data. This problem for the current values of the acceleration and reactions, when augmented by the boundary conditions described below, can be approached by a direct application of the ideas described in § 2. The required correspondences are listed in column (g) of table 1, from which it is clear that (5.12) and (5.13) are the present form of (2.8) and (2.9) respectively. The governing Hessian in (2.28) is just ρ times the metric tensor g_{jk}, from which it is clear that the quadratic functions X and Y are convex.

Comparing table 1(g) with (2.34) to (2.36) suggests that in place of the bilateral conditions (5.12) one might envisage unilateral conditions of the type.

$$A_\alpha^{\ ij}\ddot{u}_{j,i} \geqslant s_\alpha, \tag{5.15}$$

$$p_\alpha \geqslant 0, \tag{5.16}$$

$$p_\alpha[A_\alpha^{\ ij}\ddot{u}_{j,i} - s_\alpha] = 0. \tag{5.17}$$

The uNx of table 1(g) suggests that the Nx of (2.21) or (2.37) be taken here as the $m \times 1$ matrix $-A_\alpha^{\ ij}n_i\ddot{u}_j$. Associated unilateral boundary conditions on a part σ_c of σ would then be

$$A_\alpha^{\ ij}n_i\ddot{u}_j \leqslant -T_\alpha, \tag{5.18}$$

$$p_\alpha \geqslant h_\alpha, \tag{5.19}$$

$$(h_\alpha - p_\alpha)(T_\alpha + A_\alpha^{\ ij}n_i\ddot{u}_j) = 0, \tag{5.20}$$

where each T_α and each h_α is an assigned distribution of scalars on σ_c. Corresponding bilateral boundary conditions would be of the type

$$p_\alpha = h_\alpha \quad \text{on a part} \quad \sigma_u \quad \text{of} \quad \sigma, \tag{5.21}$$

$$-A_\alpha^{\ ij}n_i\ddot{u}_j = T_\alpha \quad \text{on a part} \quad \sigma_x \quad \text{of} \quad \sigma. \tag{5.22}$$

Here h_α and T_α are assigned distributions on σ_u and σ_x respectively, and $\sigma_u+\sigma_x+\sigma_c=\sigma$ to be definite.

We may now employ the uniqueness theorems of § 2 (v) to prove that the problem defined by (5.13) with (5.12) or (5.15) to (5.17) in τ, subject to boundary conditions such as (5.18) to (5.22) on σ, has a unique solution for the acceleration $\ddot{\boldsymbol{u}}$. Moreover, the extremum principles (2.42) and (2.44) hold, and these extremum principles are complementary provided

$$\int h_\alpha T_\alpha \,\mathrm{d}\sigma_c \leqslant 0 \tag{5.23}$$

by (2.46). The principle (2.42) minimizes

$$\int [\tfrac{1}{2}\rho \ddot{\boldsymbol{u}}^*.\ddot{\boldsymbol{u}}^* - \boldsymbol{z}.\ddot{\boldsymbol{u}}^*]\,\mathrm{d}\tau + \int h_\alpha A_\alpha{}^{ij} n_i \ddot{u}_j^* \,\mathrm{d}(\sigma_u+\sigma_c), \tag{5.24}$$

and it may be regarded as a generalization of the 'principle of extreme compulsion' (see Truesdell & Toupin 1960, § 237) to include unilateral conditions. The dual principle (2.44) maximizes

$$\int\left\{s_\alpha p_\alpha^\dagger - \frac{1}{2\rho}[z^j-(p_\alpha^\dagger A_\alpha{}^{ij})_{,i}]\,[z^k-(p_\alpha^\dagger A_\alpha{}^{ik})_{,i}]\,g_{jk}\right\}\mathrm{d}\tau + \int T_\alpha p_\alpha^\dagger \,\mathrm{d}(\sigma_x+\sigma_c). \tag{5.25}$$

As in the case of (4.10) the maximum can be improved because it is a non-homogeneous quadratic in the $p_\alpha^\dagger$.

6. Example: incompressible Newtonian fluid

For an example of § 5, suppose the only constraint is that of incompressibility. Then $m=1$, we can drop the α subscripts, and (5.4) is the zero difference between the values of the scalar triple product $[(\partial \boldsymbol{R}/\partial\theta^1)\ (\partial \boldsymbol{R}/\partial\theta^2)\ (\partial \boldsymbol{R}/\partial\theta^3)]$ in the current and the reference configurations. Suppose further that we pose the initial value problem in the reference configuration itself. There is then no distinction between nominal stress and the symmetric true stress, and we can without loss of generality introduce the further simplication that the material coordinates θ^i coincide with rectangular Cartesian spatial coordinates at this instant. The expression of (5.3) at this instant is then merely the equation of continuity

$$\dot{u}_{j,j} = 0. \tag{6.1}$$

If the material is a Newtonian viscous fluid its constitutive equations now take the familiar form

$$s_{ij} = 2\mu\epsilon_{ij} - p\,\delta_{ij}, \tag{6.2}$$

which is an example of (5.10) in which p is the mechanical pressure, μ the viscosity and

$$\epsilon_{ij} = \tfrac{1}{2}(\dot{u}_{i,j}+\dot{u}_{j,i}) \quad \text{the strain-rate, so that} \quad r_{ij} = \mu(\dot{u}_{i,j}+\dot{u}_{j,i}).$$

The incompressibility constraint is smooth in the sense specified by (5.11).

The conditions (5.12) and (5.13) now take the form

$$\ddot{u}_{j,j} = s \quad \text{where} \quad s = \dot{u}_{k,i}\dot{u}_{i,k}, \tag{6.3}$$

$$\rho\ddot{u}_j - z_j = -p_{,j} \quad \text{where} \quad z_j = \mu\dot{u}_{j,ii} + \rho b_j. \tag{6.4}$$

We recall that the current configuration is supposed known and is in this example being used as the reference configuration, and that the velocity distribution $\dot{\boldsymbol{u}}$ is supposed known and to satisfy (6.1). These facts have been used in arriving at the stated forms of s and z_j.

We may regard (6.3) as a requirement that (6.1) shall be satisfied not only at the instant $t=0$ in question but also at the immediately following instant $t=0+$. In terms of time-derivatives evaluated at a given place the acceleration has the property

$$\ddot{\boldsymbol{u}} = \frac{\partial \dot{\boldsymbol{u}}}{\partial t} + \dot{\boldsymbol{u}}.\nabla\dot{\boldsymbol{u}}. \tag{6.5}$$

Therefore, since $\dot{\boldsymbol{u}}$ is known and satisfies $\nabla.\dot{\boldsymbol{u}} = 0$, the initial value problem for $\ddot{\boldsymbol{u}}$ could alternatively be posed as a problem for $\partial\dot{\boldsymbol{u}}/\partial t$, in which (6.3) would be replaced by

$$\nabla.[\partial\dot{\boldsymbol{u}}/\partial t] = 0. \tag{6.6}$$

Therefore, whatever the given r^{ij}, when the only constraint is incompressibility the problem only has content for unsteady flows, whose development might be followed iteratively—i.e. by reposing the problem after each of a series of time increments.

Appropriate bilateral boundary conditions would be to assign the values of p or $\boldsymbol{n}.\ddot{\boldsymbol{u}}$ at each point of the boundary σ, as (5.21) or (5.22) show. Then the initial value problem has a unique solution for $\ddot{\boldsymbol{u}}$, and ∇p is then also unique from (6.4). Hence the distribution of p is uniquely determined at the given instant if there is at least one boundary point where it is assigned. Associated complementary extremum principles may be inferred from (5.24) and (5.25).

Moreau has proposed (1964, 1966*b*, 1967), in the context of an incompressible perfect fluid ($\mu = 0$ in (6.2)), that a definition of the onset of cavitation within the fluid can be given by replacing (6.3) by the corresponding unilateral form indicated by (5.15) to (5.17), namely

$$\ddot{u}_{j,j} \geqslant s, \tag{6.7}$$

$$p \geqslant 0, \tag{6.8}$$

$$p[\ddot{u}_{j,j} - s] = 0. \tag{6.9}$$

Here p is to be regarded as the excess of the actual pressure over a known vaporization pressure, the gradient of the latter providing an extra term in the known quantity $\boldsymbol{z}$ in (6.4). Then cavitation can only begin (*in*equality in (6.7)) within the fluid if $p = 0$ in (6.8), and not if $p > 0$, as (6.9) shows. Associated unilateral boundary conditions on a part σ_c of the boundary of the fluid are given by (5.18) to (5.20) as

$$\boldsymbol{n}.\ddot{\boldsymbol{u}} \leqslant -T, \tag{6.10}$$

$$p \geqslant h, \tag{6.11}$$

$$(h-p)(T+\boldsymbol{n}.\ddot{\boldsymbol{u}}) = 0. \tag{6.12}$$

Here T is interpretable as an assigned inward acceleration of the retaining wall, and with h set equal to zero (since p is the *excess* pressure), (6.12) ensures that a cavity cannot appear at the wall unless the vaporization pressure is attained there. The previous uniqueness theorem and complementary extremum principles still apply for suitable combinations of the previous bilateral conditions with the unilateral conditions (6.7) to (6.12), as § 2 shows. That the extremum principles are complementary follows because (5.23) is satisfied with $h = 0$.

Moreau obtains the uniqueness theorem, extremum principles and certain other results for the perfect fluid ($\mu = 0$) with unilateral constraints of type (6.7) to (6.12), but from a rather different viewpoint. Moreau makes the point that if the solution of a problem posed with purely bilateral constraints is found to violate (6.8) in a certain region, this is no reason to conclude that cavitation (as he defines it) must begin through that region. The reason is simply that the problem posed with unilateral constraints is a different mathematical problem. He illustrates this point by a simple finite-dimensional initial motion example (1967).

References

Abadie, J. 1967 (Ed.) *Nonlinear programming*. Amsterdam. North-Holland.
Arthurs, A. M. 1967 *Proc. Roy. Soc. Lond.* A **298**, 97.
Canon, M., Cullum, C. & Polak, E. 1966 *J. S.I.A.M. Control* **4**, 528.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Cap, C. & Müller, G. 1967 *Z. angew. Math. Phys.* **18**, 672.
Chan, H. S. Y. 1968 *Int. J. Solids Struct.* **4**, 885.
Charnes, A., Lemke, C. E. & Zienkiewicz, O. C. 1959 *Proc. Roy. Soc. Lond.* A **251**, 110.
Collins, I. F. 1968 *J. Mech. Phys. Solids* **16**, 73.
Ericksen, J. L. 1960 *Handb. Phys.* III/**1**, 794.
Fiszdon, W. 1962 *Z. angew. Math. Mech.* **42**, T134.
Gladwell, G. M. L. 1965 5e Congrès International d'Acoustique, Liège, Paper L33.
Gladwell, G. M. L. & Zimmermann, G. 1966 *J. Sound Vib.* **3**, 233.
Hemp, W. S. & Chan, H. S. Y. 1966 *Oxford Univ. Eng. Sci. Dept. Rep.* 1015.66. (Optimum design of pin-jointed frameworks.)
Hill, R. 1951 *Phil. Mag.* **42**, 868.
Hill, R. 1956 *J. Mech. Phys. Solids* **5**, 66.
Hill. R. 1957*a* *J. Mech. Phys. Solids* **5**, 229.
Hill, R. 1957*b* *J. Mech. Phys. Solids* **5**, 153.
Hill, R. 1957*c* *J. Mech. Phys. Solids* **5**, 302.
Hill, R. 1959 *J. Mech. Phys. Solids* **7**, 209.
Hill, R. 1961*a* *Problems of continuum mechanics*, pp. 155–164 (S.I.A.M. Philadelphia).
Hill, R. 1961*b* *J. Mech. Phys. Solids* **9**, 114.
Hill, R. 1962 *J. Mech. Phys. Solids* **10**, 185.
Hill, R. 1963*a* *J. Mech. Phys. Solids* **11**, 305.
Hill, R. 1963*b* *W. Prager 60th Anniversary volume*. New York: Macmillan and Co.
Hill, R. 1963*c* *J. Mech. Phys. Solids* **11**, 357.
Hill, R. 1966*a* *J. Mech. Phys. Solids* **14**, 239.
Hill, R. 1966*b* *J. Mech. Phys. Solids* **14**, 95.
Hill, R. 1967 *J. Mech. Phys. Solids* **15**, 371.
Hill, R. & Power, G. 1956 *Q. Jl. Mech. appl. Math.* **9**, 313.
Hill, R. & Sewell, M. J. 1962 *J. Mech. Phys. Solids* **10**, 285.
Huang, N. C. 1965 *Tech. Rep. Stanford Univ. Div. of Engng. Mech.* no. 154.
Krajewski, B. 1963 *Archwm. Mech. stosow.* **15**, 743.
Lanczos, C. 1961 *Linear differential operators*. London: van Nostrand.
Levinson, M. 1965*a* *J. Franklin Inst.* **280**, 325.
Levinson, M. 1965*b* *J. Appl. Mech.* (September).
Luke, J. C. 1967 *J. Fluid Mech.* **27**, 395.
Lush, P. E. & Cherry, T. 1956 *Q. Jl. Mech. appl. Math.* **9**, 6.
Masur, E. F. 1954 *Q. appl. Math.* **11**, 385 (see also **14**, 315).
Masur, E. F. 1958 *Proc. U.S. natn. Congr. appl. Mech.*, p. 411.
Masur, E. F. 1968 Lecture at University College, London (June).
Moiseev, N. N. & Petrov, A. A. 1966 *Adv. appl. Mech.* **9**, 91.
Moreau, J. J. 1964 *C.r. hebd. Séanc. Acad. Sci., Paris* **259**, 3948 (gr. 2).
Moreau, J. J. 1966*a* *J. S.I.A.M. Control* **4**, 153.
Moreau, J. J. 1966*b* *J. Mech.* **5**, 439.
Moreau, J. J. 1967 *Nonlinear programming* (ed. J. Abadie), ch. 9.
Prager, W. & Taylor, J. E. 1968 *J. appl. Mech.* (March), p. 102.
Pomraning, G. C. 1968 *J. Math. Phys.* **47**, 155.
Pope, G. G. 1968 *Int. J. Mech. Sci.* **10**, 251.
Schecter, R. S. 1966 *Non-equilibrium thermodynamics, variational techniques and stability* (ed. R. J. Donnelly *et al.*), *pp.* 55–69. Chicago University Press.
Seliger, R. L. & Whitham, G. B. 1968 *Proc. Roy. Soc. Lond.* A **305**, 1.
Serrin, J. 1959 *Handb. Phys.* VIII/**1**, 125.
Sewell, M. J. 1963*a* *J. Math. Mech.* **12**, 495.
Sewell, M. J. 1963*b* *J. Mech. Phys. Solids* **11**, 377.
Sewell, M. J. 1964 *J. Mech. Phys. Solids* **12**, 279.
Sewell, M. J. 1965 *Proc. Roy. Soc. Lond.* A **286**, 402.
Sewell, M. J. 1967 *Arch. Ratl. Mechs. Anal.* **23**, 327.
Sewell, M. J. 1968*a* *Proc. Roy. Soc. Lond.* A **306**, 201.
Sewell, M. J. 1968*b* *Proc. Roy. Soc. Lond.* A **306**, 225.
Synge, J. L. 1957 *The hypercircle in mathematical physics*. Cambridge University Press.
Truesdell, C. & Toupin R. A. 1960 *Handb. Phys.* III/**1**, 226.
Truesdell, C. & Toupin, R. A. 1963 *Arch. Ratl. Mechs. Anal.* **12**, 1.
Truesdell, C. & Noll, W. 1965 *Handb. Phys.* III/**3**.
Whitham, G. B. 1967 *Proc. Roy. Soc. Lond.* A **299**, 6.
Ziegler, H. 1963 *Progress in solid mechanics* IV, chap. 2. Amsterdam: North-Holland.